

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 09:09:01 ; Search time 6808 Seconds
(without alignments)
5676.280 Million cell updates/sec

Title: US-09-617-099B-1
Perfect score: 8285
Sequence: 1 MSAPLGRGRPATPAASQP.....TTRASOSSLESSTGSPYSRS 1590

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Zgapop 6.0, Zgapext 7.0
Delep 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09617099/runat_14112003_191546_23691/app_query.fasta_1.1735
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09617099@cgn_1_1_5271@runat_14112003_191546_23691 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estl:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_estfun:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1622	19.6	1346	11	AK0831172	AK0831172 Mus muscu
2	1522	18.4	1200	11	EC018057	EC018057 Homo sapi
3	1416	17.1	1073	13	BU502896	BU502896 AGENCOURT
4	1240	15.0	944	13	BU140296	BU140296 AGENCOURT
5	1224.5	14.8	944	13	AK032619	AK032619 Mus muscu
6	1157	14.0	698	14	CB527998	CB527998 UI-M-FYO-
7	1155	13.9	799	14	CB519497	CB519497 UI-M-GHO-
8	1149	13.9	686	10	BB428808	BB428808 BB428808
9	1085.5	13.1	915	13	BU413561	BU413561 602934853
10	1048.5	12.7	1004	14	CD301064	CD301064 AGENCOURT
11	1039	12.5	1718	11	AK043152	AK043152 Mus muscu
12	1037	12.5	697	13	BU741736	BU741736 UI-E-EOL-
13	1006.5	12.1	1838	11	BC016274	BC016274 Mus muscu
14	1006.5	12.1	3080	11	BC037619	BC037619 Mus muscu
15	999	12.1	663	14	CA346437	CA346437 677333 NC
16	972.5	11.7	934	13	BU930896	BU930896 AGENCOURT
17	940.5	11.4	1684	11	BC024828	BC024828 Mus muscu
18	934.5	11.3	751	14	CB244903	CB244903 UI-M-FYO-
19	928.5	11.2	747	12	BI732490	BI732490 603353484
20	926	11.2	609	12	BM072612	BM072612 fv09007.Y
21	924.5	11.2	877	12	BI872262	BI872262 603396602
22	921.5	11.1	740	10	EG04784	EG04784 602420976
23	911	11.0	697	14	CD237320	CD237320 FNPJAE04
24	908	11.0	603	2	HS0067824	HS0067824 Homo sapi
25	901.5	10.9	740	12	BI730821	BI730821 603351819
26	868	10.5	540	13	BU338952	BU338952 603513503
27	868	10.5	657	14	BU722517	BU722517 BX722517
28	863	10.4	762	10	EG675755	EG675755 602621981
29	855.5	10.4	1001	14	CA966376	CA966376 CCLX05a16
30	845	10.2	577	13	BU211560	BU211560 604159027
31	842	10.2	567	13	BU728456	BU728456 UI-E-CL1-
32	828	10.0	731	14	CB524260	CB524260 UI-M-FYO-
33	823.5	9.9	932	12	BM035114	BM035114 fu38804.Y
34	823	9.9	557	9	AL519002	AL519002 AL519002
35	820.5	9.9	1071	13	EX345439	EX345439 EX345439
36	820	9.9	939	13	EX331506	EX331506 EX331506
37	816.5	9.9	964	13	BU523891	BU523891 AGENCOURT
38	806	9.7	590	9	AA476826	AA476826 zw95901.r
39	803	9.7	537	13	EX104117	EX104117 BX104117
40	787	9.5	542	12	BI844762	BI844762 fcl7402.Y
41	784	9.5	474	9	AA772858	AA772858 ab70a10.r
42	782.5	9.4	754	10	EG207226	EG207226 RST26694
43	781	9.4	571	10	BE750318	BE750318 201646 MA
44	776	9.4	523	9	AA602733	AA602733 np17904.s
45	768	9.3	502	10	BE501603	BE501603 ht60g10.X

ALIGNMENTS

RESULT 1
AK0831172

LOCUS
DEFINITION

AK0831172
Mus musculus adult male hippocampus cDNA, RIKEN full-length

enriched library, clone:C630023J21 product:Rab3 interacting protein

1, full insert sequence.

ACCESSION
AK0831172

VERSION
AK0831172.1

KEYWORDS
HTC; CAP trapper.

SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus

AK0831172
Mus musculus adult male hippocampus cDNA, RIKEN full-length
enriched library, clone:C630023J21 product:Rab3 interacting protein

1, full insert sequence.
AK0831172
AK0831172.1
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	1	Carninci, P. and Hayashizaki, Y.	Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
AUTHORS			
TITLE		High-efficiency full-length cDNA cloning	
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)	
MEDLINE		99279253	
PUBMED		10349636	
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
AUTHORS			Please visit our web site for further details.
TITLE		Normalizaiton and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	URL: http://genome.gsc.riken.go.jp/
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)	URL: http://fantom.gsc.riken.go.jp/.
MEDLINE		20493974	
PUBMED		11042159	
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	Location/Qualifiers
AUTHORS			1. .1346
TITLE		RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	/organism="Mus musculus"
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)	/mol_type="mRNA"
MEDLINE		20530913	/strain="CS7BL/6J"
PUBMED		11076861	/db_xref="FANTOM DB:C630023J21"
REFERENCE	4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Harai, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Tomita, M., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boifelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seiya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boxis, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.	/dev_stage="adult"
AUTHORS			290. .>1345
TITLE		Functional annotation of a full-length mouse cDNA collection	/note="unnamed protein product; Rab3 interacting protein 1 (MGD MG1:2152972, GB NM_053271, evidence: BLASTN, 100%, match=893)"
JOURNAL		Nature 409 (6821), 685-690 (2001)	/codon_start=1
MEDLINE		21085660	/protein_id="BAC38794.1"
PUBMED		11217851	/db_xref="GI:26350309"
REFERENCE	5	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	/translation="MQFTLRQVNSVLSPHGVFSSPPNLLQNELFGOTLNNARKRS PSVSDQNERVQSEEDYQVPSDCTMPERSDSDYADRSDRQRPQYEEPCHLNLR DSNRRGHRSHKEYIYDDSDVSDYERQREEFYQARYRSDFNARIPVKKQPIEEQ MRHAEVSRARHRRHSDVSLANAEISRLMRDRSPRSVSRRAAMENQRSYK TNERTEAGQSSYQPTNSHPPTPRSPILDRPDMRRADSLRKHHLDPSSVRK TKRRTMETLRNDSLSQSSQSVRPPPPRPHKSKGKMRQVLSLSSEELASTPEYT SCDDVLESESVSEKGS"
AUTHORS			BASE COUNT 443 a 275 c 334 g 294 t
TITLE		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	ORIGIN
JOURNAL		Nature 420, 563-573 (2002)	Alignment Scores:
MEDLINE		11217851	Pred. No.: 4,05e-103 Length: 1346
PUBMED			Score: 1622.00 Matches: 312
REFERENCE	6	(bases 1 to 1346)	Percent Similarity: 100.00% Conservative: 0
AUTHORS			Best Local Similarity: 100.00% Mismatches: 0
TITLE		Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Harai, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, K., Saiton, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	Query Match: 19.58% Indels: 0
AUTHORS			DB: 11 Gaps: 0
TITLE		Direct Submission	US-09-617-099B-1 (1-1590) x AK083172 (1-1346)
JOURNAL		Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome	QY 273 ArgLysArgSerProSerValSerArgAspGlnAsnArgArgTyrGluGlnSerGluGlu 292
MEDLINE			DB 410 AGAAAACGAAGTCCATCAGTGTCCAGGATCAAAATCGAAGATACGAGCAAGTGAAGAA 469
PUBMED			QY 293 ArgGluAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAsp 312
REFERENCE			DB 470 AGAGAGGACTACTCAGATATGTTCTTCAGATGTGTACATGCCAAGATCTCTTCGGAT 529
AUTHORS			QY 313 TyrAlaAspArgArgSerGlnArgGluProGlnPheTyrGluGluProGlyHisLeuAen 332
TITLE			DB 530 TATGCTGATAGCATCTCAGGTGAGCTCAATTTTATGAGAACCTGGTCAATTAAT 589
JOURNAL			QY 333 TyrArgAspSerAsnArgArgGlyHisArgHisSerLysGluTyrIleValAspAspGlu 352
MEDLINE			DB 590 TACAGGATTCCTAACAGGAGAGCCATAGACATTCCTCAAGAGTATATTGTCGATGATGAA 649
PUBMED			QY 353 AspValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGlnAlaArg 372
REFERENCE			DB 650 GATGTGGAGAGCAGAGATGAATATGAAACACAAAGAGAGAGAGGAATACCAGGACCGC 709

QY 373 TyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGln 392
 Db 710 TACAGAGTGTATCCAAATCTGCGCGGTATCCCGTAAAGCCACCACTACGAGAACAA 769

QY 393 MetArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSer 412
 Db 770 ATGCGATCCAGCTGAGGTCTCAGGCGACGACATGAGAGGACGACAGTGAITTTCT 829

QY 413 LeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspArgProSer 432
 Db 830 TTGGCAACCGCTGAACATAGAGATTCAGAGTTCCTCTGCTAAGATGATAGACCATCA 899

QY 433 ArgGlnArgSerValSerGlnArgAlaAlaMetGlnArgGlnArgSerTyrSerMet 452
 Db 890 AGGCAAGATCTGTATCTGACGCTAGAGTCAATGGAACCAACCAACGATCGTATTCATG 949

QY 453 GluArgThrArgGluAlaGlnGlyGlnSerSerTyrProGlnArgThrSerAsnHisSer 472
 Db 950 GAAAGAACTCGAGAGGCTCAGGACAAAGTTCCTATCCACAAAGGACCTCAATCATAGT 1009

QY 473 ProProThrProArgValSerProIleProLeuAspArgProAspMetArgAlaAsp 492
 Db 1010 CTTCCACCCCTCGCGGAGCGCTATACCGCTTGATAGACCAACATGAGCGCGCTGAC 1069

QY 493 SerLeuArgLysGlnHisLeuAspProSerSerAlaValArgLysThrLysArgGlu 512
 Db 1070 TCCCTACGAAACAGCACCTTAGATCCAGCTCTGCTGAGAGAAACGAGCGAGAA 1129

QY 513 LysMetGluThrMetLeuArgAsnAspSerLeuSerSerAspGlnSerGluSerValArg 532
 Db 1130 AAAATGGAAACCATGTTAAGGAATGATTTCTTTGAGTTTCAGACCACTCGAGTCAAGT 1189

QY 533 ProProProProArgProHisLysSerLysLysGlyLysMetArgGlnValSerLeu 552
 Db 1190 CGCGCCCCCACCAGGCGCTCAATAATCCAAAGAAAGAGGATAAATGCGCCAGGTTTCACTG 1249

QY 553 SerSerSerGluGluGluLeuAlaSerThrProGluTyrThrSerCysAspAspValGlu 572
 Db 1250 AGCAGCTCGAGAGGAGGAGCTGGCATCCACACTGAGTATACAACTGTGATGATGAG 1309

QY 573 LeuGluSerGluSerValSerGluLysGlyAspSer 584
 Db 1310 CTGGAAGCGAGGTGTGAGTGAGAAAGGGGACAGT 1345

RESULT 2
 BC018057
 LOCUS
 DEFINITION Homo sapiens, Similar to KIAA0751 gene product, clone
 IMAGE:4792768, mRNA.
 ACCESSION BC018057
 VERSION BC018057.1 GI:17390117
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1200)
 Strausberg, R.
 Direct Submission
 Submitted (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapps-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 32 Row: b Column: 12
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 7662285
 This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers

1..1200
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4792768"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

BASE COUNT 431 a 220 c 283 g 266 t
 ORIGIN

Alignment Scores:

Pred. No.: 3 38e-96 Length: 1200
 Score: 1522.00 Matches: 288
 Percent Similarity: 95.60% Conservative: 16
 Best Local Similarity: 90.57% Mismatches: 14
 Query Match: 18.37% Indels: 0
 DB: 11 Gaps: 0

US-09-617-099b-1 (1-1590) x BC018057 (1-1200)

QY 273 ArgLysArgSerProSerValSerArgAspGlnAsnArgArgTyrGluGlnSerGluGlu 292
 Db 242 AGGAAAGAAAGGCCATCTGTCTCCAGATCAGATAGATAGATACGACCAAGGAGAA 301

QY 293 ArgGluAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAsp 312
 Db 302 AGAGAGGAATATTCAGATGCTCTTCGGATACCGCAATGCGCTAGATCTCCATCATGAT 361

QY 313 TyrAlaAspArgArgSerGlnArgGluProGlnPheTyrGluGluProGlyHisLeuAsn 332
 Db 362 TATGCTGATAGCGGATTTCAACATGACCTCAGTTTATGAGACTCTGATCATTTAAGT 421

QY 333 TyrArgAspSerAsnArgArgGlyHisArgHisSerLysGluTyrIleValAspAspGlu 352
 Db 422 TATAGGACCTCCAAAGGAGAGTCTAGGCAATTCCTCAAGAGATATATTTGATAGATCATG 481

QY 353 AspValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGlnAlaArg 372
 Db 482 GATGTGGAAAGCAGAGATGATATCGAAAGGCAAGGAGAGAGAGATACCGTCCAGC 541

QY 373 TyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGln 392
 Db 542 TACCGAAGTGTATCCGAAATTTGCGCGGTATTCAGTAAAGCCACCAACCTATGAAGACAA 601

QY 393 MetArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSer 412
 Db 602 ATGCGGATCCATGCTGGAAGTGTCCGACGACGCGATGAGAGAGGATAGTGTATTTCT 661

QY 413 LeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspArgProSer 432
 Db 662 TTGGCAACCGCTGAACATAGAGATTCAGAGATTTCTATGCTAAGATGATGATCGACCATCA 721

QY 433 ArgGlnArgSerValSerGluArgAlaAlaMetGlnAsnArgGlnArgSerTyrSerMet 452
 Db 722 AGGCAAGATCTATATCATGAGAGTGTGCGCATGGAATATCAGCGATCTTATTCATG 781

QY 453 GluArgThrArgGluAlaGlnGlyGlnSerSerTyrProGlnArgThrSerAsnHisSer 472
 Db 782 GAAAGAACTCGAGAGGCTCAGGACCAAGTCTTTATGACCAAGGACCAACCAACCATAGT 841

473 ProProThrProArgSerProileProLeuAspArgProAspMetArgAlaAsp 492
 842 CTTCTACCCAGGAGGAGTCCACTCCATAGATAGACAGGAGTCTGAGGCTACTGAC 901
 493 SerLeuArgGluHisLeuAspSerSerSerAlaValArgLysThrLysArgGlu 512
 902 TCATACGGAACACACACACCTAGATCTTACGCTCTGCTGTGAAGAAAAACAAACGGGAA 961
 513 LysMetGluThrMetLeuArgAsnAspSerLeuSerSerAspGlnSerValArg 532
 962 AAAATGGAAACAAATGTAAGGAATGATTTCTCTCAGTTCAGACCAAGTCAGTGTGAGA 1021
 533 ProProProArgProHisLysSerLysGlyGlyLysMetArgGlnValSerLeu 552
 1022 CTTCCACCAACCAAGGCTCATAAATCAAGAAAGCGGTAAATGCCCGCAGATTTGTTG 1081
 553 SerSerSerGluGluGluLeuAlaSerThrProGluTyrThrSerCysAspValGlu 572
 1082 AGCAGTTTCAGAGGAGGAAATGGCTTCCACGCTGAAATATACAACTGTGTGATGTTGAG 1141
 573 LeuGluSerGluSerValSerGluLysGlyAspSerGlnLysArgLys 590
 1142 ATTGAAGTCAGAGTGTAAAGTAAAGGAGGACAGTCAGTCAAGGAAAAAAGAAAAA 1195

RESULT 3

LOCUS BU502896 1073 bp mRNA linear EST 12-SEP-2002
 DEFINITION AGENCOURT_8929724 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6489689
 5', mRNA sequence.

ACCESSION BU502896
 VERSION BU502896.1 GI:22808299

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 1073)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: The Cepko Laboratory

CDNA Library Preparation: Life Technologies, Inc.

CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Clone Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM14038 row: e column: 18

High quality sequence stop: 571.

FEATURES

source

Location/Qualifiers
 1..1073
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6489689"
 /tissue_type="retina"
 /lab_host="PH108 (phage-resistant)"
 /clone_lib="NIH_MGC_94"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dr primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 340 a 238 c 277 g 217 t 1 others

ORIGIN

Alignment Scores:
 Pred. No.: 7,44e-89 Length: 1073
 Score: 1416.00 Matches: 320
 Percent Similarity: 75.81% Conservative: 9
 Best Local Similarity: 73.73% Mismatches: 19

Query Match: 17.09% Indels: 86
 DB: 13 Gaps: 4

US-09-617-099b-1 (1-1590) x BU502896 (1-1073)

354 ValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGlnAlaArgTyr 373
 DB 1 GTGAGAGCAGAGATGAATATGAAGACAAAGAGAGAGAGAGGAGATATCCAGGCACGCTAC 60
 374 ArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMet 393
 DB 61 AGAAGTGATCCAAATCTGCCCGGTATCCCGTAAAGCCACCAACCTACGAGAACAAATG 120
 394 ArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSerLeu 413
 DB 121 CGCATCCACGCTGAGGTGTCCAGGCGCACCATAGAGAGAGGACAGTGTATTTCTTTG 180
 414 AlaAsnAlaGluLeuAspSerArgIleSerLeuLeuArgMetAspArgProSerArg 433
 DB 181 GCAAAACGCTCAACTAGAGATTCAGGATTTCTCTGCTAAGGATGATAGACCATCAAGG 240
 434 GlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGlu 453
 DB 241 CAAAGATCTGTATCTGAACGCTAGAGCTGCAATGGAACCAACCAACGATCGTATTCAATGAA 300
 454 ArgThrArgGluAlaGlnGlnSerSerTyrProGlnArgThrSerAsnHisSerPro 473
 DB 301 AGAATCCGAGAGGCTCAGGACCAAGTCTTATCCCAAGGACCTCAATCATAGTCTCT 360
 474 ProThrProArgSerProIleProLeuAspArgProAspMetArgArgAlaAspSer 493
 DB 361 CCCACCCCTCGCGGAGCCCTATACGCTGTATAGACCAAGATGAGGCGCGCTGACTCC 420
 494 LeuArgLysGlnHisHisLeuAspProSerSerAlaValArgLysThrLysArgLys 513
 DB 421 CTACGGAACACAGCACCACTTAGATCCAGCTCTGCTGTGAGAAACGAAGCGAGAAAA 480
 514 MetGluThrMetLeuArgAsnAspSerLeuSerSerAspGlnSerGluSerValArgPro 533
 DB 481 ATGAAACCACTGTTAAGGAAATGATTTCTTGTAGTTCAGACCACTGCGAGTCAGTGGCGG 540
 534 ProProProArgProHisLysSerLysGlyGlyLysMetArgGlnValSerLeuSer 553
 DB 541 CCCCAACCAAGGCTCATTAATCCAGAAAGAGGAGTAANAATGCGCAGGTTTCACTGAGC 600
 554 SerSerGluGluGluLeuAlaSerThrProGluTyrThrSerCysAspAspValGlu 573
 DB 601 AGCTCGGAGGAGGAGGCTGCATCCACCTGAGTATACAAAGCTGTGTATGATGTGAGCT 660
 573 uGluSerGluSerValSerGluLysGlyAspSerGlnLysGlyLysArgLysThrSerG1 593
 DB 661 GGAAAGCGAGAGTGTGAGTGAAG----- 682
 593 uGlnGlyValLeuSerAspSerAsnThrArgSerGluArgGlnLysLysArgMetTyrTy 613
 DB 682 ----- 682
 613 rGlyGlyHisSerLeuGluAspLeuGluTtpSerGluProGlnIleLysAspSerG1 633
 DB 682 ----- 682
 633 yValAspThrCysSerSerThrThrLeuAsnGluGluHisSerHisSerAspLys-HisP 653
 DB 683 ----- 683
 653 roValThrTtpGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuAsnL 673
 DB 691 CTGTACCTGCGACCCATCCAAAGATGGAGATCGCTTAATTTGGCGGTATTTTATTAAATA 750
 673 ysArgLeuLysAsp-GlySerValProArgAspSerGlyValaMetLeuGlyLeuLysVal 692
 DB 751 AGCGTTTTAAAGATGGGAAGTGTACCTCGAGACTCGAGAGCAATGCTGGCGCTTAAGGT 810

693	QY	Val -GlyGlyLeuMetThrGluSer -GlyArgLeuCysAlaPheIleThrLysValLysL	712	
811	Db	811	GTAAGAAGAGAAAAGAGACTGAATCCAGGCCACCTTTGGGCATTATTATTACCAAGTAAAAA	870
712	QY	ysGlySerLeuAla -AspThrVal -GlyHisLeuArgProGlyAspGluValLeu----	729	
871	Db	871	AGGAAATTTAGCTTGAAACTGTGAGACATCTTAACACAGGGGAGTAATCTTCGAAA	930
730	QY	-GluTrpAsnGlyArgLeuLeuGlnGlyAlaThrPhe -GluGluValTyrAsnIleIle-	748	
931	Db	931	GGAATGGGAGGC--TATTGCCAGGAGCCCTCTTTGGGGAAGTTTACACATTATTC	987
749	QY	LeuGluSerIysPro--GluProGlnValGluLeuValVal -SerArgProIleGlyAsp	767	
988	Db	988	CTAAATCCCAACTCGAACCCACCAATTTCATCTGTGTGTTTCCAGGCCCATGTGGAAA	1047
768	QY	768	Ile--ProArgIlePro 772	
1048	Db	1048	TATTCTCTTAATATCCCC 1065	
RESULT 4				
BUI40296				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
COMMENT				
FEATURES				
source				
<p> BUI40296 944 bp mRNA linear EST 25-NOV-2002 60313465F1 CSQCHL24 Gallus gallus cDNA clone CHEST116b17 5', mRNA sequence. BUI40296 BUI40296.1 GI:25355131 EST. Gallus gallus (chicken) Gallus gallus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus. 1 (bases 1 to 944) Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J. A Comprehensive Collection of Chicken cDNAs Curr. Biol. 12 (22), 1965-1969 (2002) 2335534 12445392 Contact: Simon Hubbard Department of Biomolecular Sciences University of Manchester Institute of Science and Technology (UMIST) PO Box 88, Manchester, M60 1QD, UK Tel: 01612008930 Fax: 01612360409 Email: Simon.Hubbard@umist.ac.uk. Location/Qualifiers 1. .944 /organism="Gallus gallus" /mol_type="mRNA" /strain="White Leghorn, Hisex" /db_xref="taxon:9031" /clone="CHEST116b17" /dev_stage="16 day embryo" /lab_host="DH10B" /clone_lib="CSEQHL24" /notes="Organ: brain; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; Modification of pBluescript II KS(+) [Stratagene] vector to accommodate cDNA produced with the T-trimmed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387 ,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BglI and BamHI sites [5'ggcccggtgcagcccgatccgaaaaaag] [5'aattcttttttgatccgggtcgacgc]" </p>				
BASE COUNT				
ORIGIN				
321 a 212 c 236 g 175 t				

Score:	1240.00	Matches:	247
Percent Similarity:	88.67%	Conservative:	27
Best Local Similarity:	79.94%	Mismatches:	32
Query Match:	14.97%	Indels:	4
DB:	13	Gaps:	0

US-09-617-099B-1 (1-1590) x BU140296 (1-944)	
Qy 222	LysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAlaVal 241
Db	3 AAGCTCCAAGAGCATTTTCAGTACCAAGGACCACACAGGTGACATATCCACACA-AGTTTG 61
Qy 242	GluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGlyVal 261
Db	62 GACAAAACACAGATCTTCAGGACTCACCAGGCAAGACTCAATTAAAGAAATGGTTCTCGAGTA 121
Qy 262	LysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSerArg 281
Db	122 AAGCACCATAATACAGGTGACACAGCATCAGACAGGAAAAGAGCCCATCAATATCAAGA 181
Qy 282	AspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrValPro 301
Db	182 GAACAGACAGAAAGATACGACCAAGGAAGAGAGACGAATATTCACAGTATGCTACA 241
Qy 302	SerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArgGlu 321
Db	242 TCAGACAGTGGCATGCCAGTCCACCATCAGATTATTCGATAGCGGTTCACATCGTGGC 301
Qy 322	ProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGlyHis 341
Db	302 CCTCAGTTATACGAAGAGCCCTGAACTGGGTGATTATAGAGATTCCACAGGAGGAGTGGC 361
Qy 342	ArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyrGlu 361
Db	362 AGCGTTCCAGAGAGTATCCGGTAGAGAGAGAGGATGCACAGAACAGGAGAGAAATATGAA 421
Qy 362	ArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAlaArg 381
Db	422 AGGCAGAGGAGAGAAGAGGAATATACGCACGATACAGAGTGCATCTAAATTTGGCTCGT 481
Qy 382	TyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValSerArg 401
Db	482 TATCCGGTCAAGCCACACCAATATCAGGAACAAATGCGTATCCATGCTGAAGTGTCCCGA 541
Qy 402	AlaArgHisGluArgArgGlyHisSerAspValSerLeuAlaAsnAlaGluLeuGluAspSer 421
Db	542 GCACGTATGAACCGAGACACAGGTATGCTCATTTGGCAAAATCTAGTGTGAAAGATTC 601
Qy 422	ArgIleSerLeuLeuMetAspArgProSerArgGlnArgSerValSerGluArgArg 441
Db	602 CGGATATCTATGCTGAGATGGAACGGCGGTCAAGGCAAAAGATCGGTGCTCAGCGCAGA 661
Qy 442	AlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGlyGln 461
Db	662 GCTGCCATGGAAAAATCAGCGTTCTACTCTATGGAAGAACTCGAGAGGCTCAGGAGCCA 721
Qy 462	SerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArg-SerProIle 481
Db	722 AGTCCCCAATCGTCAGAGACCACAAAATCATAGCCCTCTACACCTAGGAGGAAGTCCAA 781
Qy 481	eProLeuAspArgProAsp-MetArgArgAlaAspSerLeuArgLysGlnHisHisLeuAla 501
Db	782 TCCTCTGGAAGACCGGACATTTAGGCGCTCTGATTCTATTAGGAAACACACCACTTTGG 841
Qy 501	spProSerSerAlaValArgLysThrLysArgGlu-LysMetGluThrMetLeuArgAsn 520
Db	842 ATCCCAATTCCTCGGTACGGAACACGAAACCGTGAAACGAATGGAGACCATGTTACGGAAT 901
Qy 521	AspSerLeuSerSerAspGln 527
Db	902 GATTCACTCAGCTCAGACCAA 922

RESULT 5

AK032619 3687 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:6430709C16 product:Rab3 interacting protein 1, full insert sequence.

AK032619 1 GI:26328428
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

2 99279253
MEDLINE
10349636

3 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

4 20499374
MEDLINE
11042159

5 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hatada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

6 20530913
MEDLINE
11076961

7 Kawaji, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Buit, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.P., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Walming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

8 21085660
MEDLINE
11217851

9 Alignment Scores:
Pred. No.: 7.39e-75 Length: 3687
Score: 1224.50 Matches: 266
Percent Similarity: 78.47% Conservative: 22
Best Local Similarity: 72.48% Mismatches: 47
Query Match: 14.78% Indels: 33
DB: 11 Gaps: 7

10 US-09-617-099B-1 (1-1590) x AK032619 (1-3687)

11 QY 1245 ArgThrSerSerAla---SerArgPheSer-----SerThrSerTyrMetSerVal 1360
Db 9 CGGACTTCACCGTGACACGCCGTCTCTTGACAGCTTCTCCCTGACCCCTCCCCAGC 68
QY 1261 GlnSerGluArgProArgGlyAsnArgIleSerValPheThrSerIysMetGlnAsn 1280

AK032619 3687 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male olfactory brain cDNA, RIKEN full-length
enriched library, clone:6430709C16 product:Rab3 interacting protein
1, full insert sequence.

AK032619 1 GI:26328428
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
95279253
10349636

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hatada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076961

4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kushl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldairelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Buit, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-oka, K., Wang, K.H., Weltz, C., Whittaker, C., Wilm, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3687)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Ishikawa, K., Itoh, M., Kawai, J., Kawauchi, T., Kikuchi, T., Kikuchi, Y.,
Kishimoto, A., Kishimoto, T., Kishimoto, Y., Kishimoto, Z., Kishimoto, A.,
Kishimoto, B., Kishimoto, C., Kishimoto, D., Kishimoto, E., Kishimoto, F.,
Kishimoto, G., Kishimoto, H., Kishimoto, I., Kishimoto, J., Kishimoto, K.,
Kishimoto, L., Kishimoto, M., Kishimoto, N., Kishimoto, O., Kishimoto, P.,
Kishimoto, Q., Kishimoto, R., Kishimoto, S., Kishimoto, T., Kishimoto, U.,
Kishimoto, V., Kishimoto, W., Kishimoto, X., Kishimoto, Y., Kishimoto, Z.,
Kishimoto, A., Kishimoto, B., Kishimoto, C., Kishimoto, D., Kishimoto, E.,
Kishimoto, F., Kishimoto, G., Kishimoto, H., Kishimoto, I., Kishimoto, J.,
Kishimoto, K., Kishimoto, L., Kishimoto, M., Kishimoto, N., Kishimoto, O.,
Kishimoto, P., Kishimoto, Q., Kishimoto, R., Kishimoto, S., Kishimoto, T.,
Kishimoto, U., Kishimoto, V., Kishimoto, W., Kishimoto, X., Kishimoto, Y.,
Kishimoto, Z., Kishimoto, A., Kishimoto, B., Kishimoto, C., Kishimoto, D.,
Kishimoto, E., Kishimoto, F., Kishimoto, G., Kishimoto, H., Kishimoto, I.,
Kishimoto, J., Kishimoto, K., Kishimoto, L., Kishimoto, M., Kishimoto, N.,
Kishimoto, O., Kishimoto, P., Kishimoto, Q., Kishimoto, R., Kishimoto, S.,
Kishimoto, T., Kishimoto, U., Kishimoto, V., Kishimoto, W., Kishimoto, X.,
Kishimoto, Y., Kishimoto, Z., Kishimoto, A., Kishimoto, B., Kishimoto, C.,
Kishimoto, D., Kishimoto, E., Kishimoto, F., Kishimoto, G., Kishimoto, H.,
Kishimoto, I., Kishimoto, J., Kishimoto, K., Kishimoto, L., Kishimoto, M.,
Kishimoto, N., Kishimoto, O., Kishimoto, P., Kishimoto, Q., Kishimoto, R.,
Kishimoto, S., Kishimoto, T., Kishimoto, U., Kishimoto, V., Kishimoto, W.,
Kishimoto, X., Kishimoto, Y., Kishimoto, Z., Kishimoto, A., Kishimoto, B.,
Kishimoto, C., Kishimoto, D., Kishimoto, E., Kishimoto, F., Kishimoto, G.,
Kishimoto, H., Kishimoto, I., Kishimoto, J., Kishimoto, K., Kishimoto, L.,
Kishimoto, M., Kishimoto, N., Kishimoto, O., Kishimoto, P., Kishimoto, Q.,
Kishimoto, R., Kishimoto, S., Kishimoto, T., Kishimoto, U., Kishimoto, V.,
Kishimoto, W., Kishimoto, X., Kishimoto, Y., Kishimoto, Z., Kishimoto, A.,
Kishimoto, B., Kishimoto, C., Kishimoto, D., Kishimoto, E., Kishimoto, F.,
Kishimoto, G., Kishimoto, H., Kishimoto, I., Kishimoto, J., Kishimoto, K.,
Kishimoto, L., Kishimoto, M., Kishimoto, N., Kishimoto, O., Kishimoto, P.,
Kishimoto, Q., Kishimoto, R., Kishimoto, S., Kishimoto, T., Kishimoto, U.,
Kishimoto, V., Kishimoto, W., Kishimoto, X., Kishimoto, Y., Kishimoto, Z.,
Kishimoto, A., Kishimoto, B., Kishimoto, C., Kishimoto, D., Kishimoto, E.,
Kishimoto, F., Kishimoto, G., Kishimoto, H., Kishimoto, I., Kishimoto, J.,
Kishimoto, K., Kishimoto, L., Kishimoto, M., Kishimoto, N., Kishimoto, O.,
Kishimoto, P., Kishimoto, Q., Kishimoto, R., Kishimoto, S., Kishimoto, T.,
Kishimoto, U., Kishimoto, V., Kishimoto, W., Kishimoto, X., Kishimoto, Y.,
Kishimoto, Z., Kishimoto, A., Kishimoto, B., Kishimoto, C., Kishimoto, D.,
Kishimoto, E., Kishimoto, F., Kishimoto, G., Kishimoto, H., Kishimoto, I.,
Kishimoto, J., Kishimoto, K., Kishimoto, L., Kishimoto, M., Kishimoto, N.,
Kishimoto, O., Kishimoto, P., Kishimoto, Q., Kishimoto, R., Kishimoto, S.,
Kishimoto, T., Kishimoto, U., Kishimoto, V., Kishimoto, W., Kishimoto, X.,
Kishimoto, Y., Kishimoto, Z., Kishimoto, A., Kishimoto, B., Kishimoto, C.,
Kishimoto, D., Kishimoto, E., Kishimoto, F., Kishimoto, G., Kishimoto, H.,
Kishimoto, I., Kishimoto, J., Kishimoto, K., Kishimoto, L., Kishimoto, M.,
Kishimoto, N., Kishimoto, O., Kishimoto, P., Kishimoto, Q., Kishimoto, R.,
Kishimoto, S., Kishimoto, T., Kishimoto, U., Kishimoto, V., Kishimoto, W.,
Kishimoto, X., Kishimoto, Y., Kishimoto, Z., Kishimoto, A., Kishimoto, B.,
Kishimoto, C., Kishimoto, D., Kishimoto, E., Kishimoto, F., Kishimoto, G.,
Kishimoto, H., Kishimoto, I., Kishimoto, J., Kishimoto, K., Kishimoto, L.,
Kishimoto, M., Kishimoto, N., Kishimoto, O., Kishimoto, P., Kishimoto, Q.,
Kishimoto, R., Kishimoto, S., Kishimoto, T., Kishimoto, U., Kishimoto, V.,
Kishimoto, W., Kishimoto, X., Kishimoto, Y., Kishimoto, Z., Kishimoto, A.,
Kishimoto, B., Kishimoto, C., Kishimoto, D., Kishimoto, E., Kishimoto, F.,
Kishimoto, G., Kishimoto, H., Kishimoto, I., Kishimoto, J., Kishimoto, K.,
Kishimoto, L., Kishimoto, M., Kishimoto, N., Kishimoto, O., Kishimoto, P.,
Kishimoto, Q., Kishimoto, R., Kishimoto, S., Kishimoto, T., Kishimoto, U.,
Kishimoto, V., Kishimoto, W., Kishimoto, X., Kishimoto, Y., Kishimoto, Z.,
Kishimoto, A., Kishimoto, B., Kishimoto, C., Kishimoto, D., Kishimoto, E.,
Kishimoto, F., Kishimoto, G., Kishimoto, H., Kishimoto, I., Kishimoto, J.,
Kishimoto, K., Kishimoto, L., Kishimoto, M., Kishimoto, N., Kishimoto, O.,
Kishimoto, P., Kishimoto, Q., Kishimoto, R., Kishimoto, S., Kishimoto, T.,
Kishimoto, U., Kishimoto, V., Kishimoto, W., Kishimoto, X., Kishimoto, Y.,
Kishimoto, Z., Kishimoto, A., Kishimoto, B., Kishimoto, C., Kishimoto, D.,
Kishimoto, E., Kishimoto, F., Kishimoto, G., Kishimoto, H., Kishimoto, I.,
Kishimoto, J., Kishimoto, K., Kishimoto, L., Kishimoto, M., Kishimoto, N.,
Kishimoto, O., Kishimoto, P., Kishimoto, Q., Kishimoto, R., Kishimoto, S.,
Kishimoto, T., Kishimoto, U., Kishimoto, V., Kishimoto, W., Kishimoto, X.,
Kishimoto, Y., Kishimoto, Z., Kishimoto, A., Kishimoto, B., Kishimoto, C.,
Kishimoto, D., Kishimoto, E., Kishimoto, F., Kishimoto, G., Kishimoto, H.,
Kishimoto, I., Kishimoto, J., Kishimoto, K., Kishimoto, L., Kishimoto, M.,
Kishimoto, N., Kishimoto, O., Kishimoto, P., Kishimoto, Q., Kishimoto, R.,
Kishimoto, S., Kishimoto, T., Kishimoto, U., Kishimoto, V., Kishimoto, W.,
Kishimoto, X., Kishimoto, Y., Kishimoto, Z., Kishimoto, A., Kishimoto, B.,
Kishimoto, C., Kishimoto, D., Kishimoto, E., Kishimoto, F., Kishimoto, G.,
Kish

Db	69	AAAGCCAAAGTGCACGCCACAGCCGCGGAGCGAGGACCGGAACCC-----	113
Qy	1281	ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMet	1300
Db	114	-----CTAGGGTTCGCGGNAACGAGCCGTCGCGACCCGAGCGGTTTCGGTACACG	167
Qy	1301	-----CysSerLeuGluLysAsnAspGlySerGlnSerAspThr	1313
Db	168	CCGCAAAAGCCCTCGTATCTGTGACCGGG-AAA-----AGCGGTCTCTCCACC	217
Qy	1314	AlaValGlyAlaLeuGlyThrSerGly-----Lys	1323
Db	218	AAATGGCGCGGAGGCGCTAGGGGCGACGCGGTCTCGAGGGCGCTCCATGCGACGCTCT	277
Qy	1324	LysArgArgSerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSer	1343
Db	278	CAGAGCGGAGAGCCCTCTCGCTCTTTTGAGGCTCTAGCTGGCTATTTCCTGTATG	337
Qy	1344	ArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrVal	1363
Db	338	AACTCCCTCGAAGAA---GATGAAGAGAGAGGAGGTTAAAAGCTACGAGCACTGTT	394
Qy	1364	GlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAla	1383
Db	395	CAGAGAAGCAGGAGACCGGCTAGCAGTGGAGATGAGAACTGATGACCCGCCAGGCC	454
Qy	1384	SerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsnLeuIlePhe	1403
Db	455	AGCCGGAAATCCACAGATGGCAGCATGACAGCTATAGCTCGGAGAGAAATCGATCTTC	514
Qy	1404	ProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLeuGlyPro	1423
Db	515	CCTGGGTCCGCTGGCTCTCTGACAGCCAGTTCACTGATTCTCTGGATGGCTGGGCCCT	574
Qy	1424	AlaGlnLeuValGlyArgGlnThrIleuAlaThrProAlaMetGlyAspIleGlnValGly	1443
Db	575	GCTCAGCTAGTGGGACGCCAGACCTCGCTACTCTCTCAATGGGTGATCATTCAGGTGGGA	634
Qy	1444	MetMetAspLysLysGlnLeuGluValGluIleIleArgAlaArgGlyLeuValVal	1463
Db	635	ATGATGATTAAGAAGGACAGCTGGAGTAGAATCATCCGGGCGCGCGCTTGGTGTA	694
Qy	1464	LysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeuAspAsnGly	1483
Db	695	AAACAGGTTCCAAGACACACTGCCAGCACCGTATGTCGAAGTGTATCTGTTAGACAACGGA	754
Qy	1484	ValCysIleAlaLysLysThrLysValAlaArgLysThrLeuGluProLeuTyrGln	1503
Db	755	GTCTGCGATAGCAAAAAAGAAACCAAGGTGGCGAGAAAGCCCTGGAGCCCTGTACACG	814
Qy	1504	GlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleValTrpGly	1523
Db	815	CAGCTCTGTCTTCGAGAGAGAGCCCGCAGGGAGGGTGTGTACAGATCATGTCCTGGGGA	874
Qy	1524	AspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeuLeuAspGlu	1543
Db	875	GATTATGGTCTGATGGATCAAAATCCTTTATGGAGTGGGCCAGATATCTTTAGATGAA	934
Qy	1544	LeuGluLeuSerAsnMetValIleGlyTrpPheLysLeuPheProProSerSerLeuVal	1563
Db	935	CTGGAACTATCCAAACATGGTGATTGGATGGTTCAAAACCTCTCCCTCTCTCTCCCTAGTA	994
Qy	1564	AspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGluSerThr	1583
Db	995	GATCCAACTTGGCACCTCTGCACAGAGAGCTTCCCAATCGTCTCTGGAAAGTTCTTACC	1054
Qy	1584	GlyProSerTyrSerArgSer	1590
Db	1055	GGACCTTCTTACTCTCGTTCA	1075

RESULT 6

CB527998

RESULT 6
CB527998

LOCUS CB527998 698 bp mRNA linear EST 28-MAR-2003
DEFINITION UI-M-FYO-cfk-j-20-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE: 6850773 5', mRNA sequence.
ACCESSION CB527998
VERSION CB527998.1 GI:29361471
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 698)
AUTHORS NTH-WGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
(This clone was contributed by the Brain Molecular Anatomy Project
(BMAP))
Seq primer: pYX-5.
FEATURES
source Location/Qualifiers
1. 698
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6850773"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/notes="Organ: Brain; Vector: pYX-Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCAGACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
BASE COUNT 232 a 152 c 173 g 139 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 5,38e-71 Length: 698
Score: 1157.00 Matches: 228
Percent Similarity: 98.28% Conservative: 1
Best Local Similarity: 97.85% Mismatches: 3
Query Match: 13.96% Indels: 2
DB: 14 Gaps: 0
US-09-617-0998-1 (1-1590) x CB527998 (1-698)
Qy 289 GlnSerGluArgGluAspTyrSerGluTyrValProSerAspGlyThrMetProArg 308
Db 1 CAATAGTGAAGAAAGAGAGGACTACTACAGTATGTTCCTTCAGATGGTACATGCCAAGA 60
Qy 309 SerProSerAspTyrAlaAspArgArgSerGlnArgGluProGlnPheTyrGluGluPro 328
Db 61 TCTCCTTCGGATATGCTGATAGACGATCTACGCTGAGCCTCAATTTTATGAGAACCT 120
Ov 329 GlyHisLeuAsnTyrArgAspSerAsnArgGlyHisArgHisSerIysGluTyrIle 348

Db	121	GGTCATTAAATACAGGGATCTTAAACAGGAGAGCCATAGACATTCCTCAAGAGTATATT	180
Qy	349	VallAspGluSerValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGlu	368
Db	181	GTGGATGATGAGATGTGGAGAGAGATGATATGAAGACAAAGAGAGAGAGGAA	240
Qy	369	TyrGlnAlaArgTyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnPro	388
Db	241	TACCAGGCAGCTACAGAGTATCCAAATCTGGCCGGTATCCGCTAAAGCCCAACCC	300
Qy	389	TyrGluGluGlnMetArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHis	408
Db	301	TACGAAGAACAAATCGCATCCAGCTGAGTGTCCAGGCGACCATGAGAGAGGCAC	360
Qy	409	SerAspValSerLeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMet	428
Db	361	AGTATGATGTTTCTTGGCAACGCTGAATAGAGATCCAGGATTTCTCTCTAAGGATG	420
Qy	429	AspArgProSerArgGlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArg	448
Db	421	GATAGACCATCAAGCAAGATCTGATCTGAACGTAGAGCTGCAATCGAAACCAACGA	480
Qy	449	SerTyrSerMetGluArgThrArgGluAlaGlnGlyGlnSerSerTyrProGlnArgThr	468
Db	481	TCGTATTCAATGGAAGAACTCGAGAGGCTCAGGAGCAAAAGTTCTTATCCAAAGGACC	540
Qy	469	SerAsnHisSerProProThrProArgArgSerProIleProLeuAspArgProAspWe	488
Db	541	TCAATCATATGCTCTCCACCCCTCGGCGAGCCCTATACCGCTTGATAGACACAT	600
Qy	488	LArgArgAlaAspSerLeuArgGlnHisHisLeuAspProSerSerAlaValArgLys	508
Db	601	GAGGCGCGTACTCCCTACGGAACAGAC-CACITAGATCCAGCTCTGCTGTGAGGA	659
Qy	508	sThrylsArgGluLysMetGluThrMetLeuArgAsn	520
Db	660	AACGAGCGAGAAAAATGGAACCATGNTAGGAAT	696
RESULT 7			
CB519497			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
Seq primer: pYX-5.			
Location/Qualifiers			
1..799			
/organism="Mus musculus"			
/mol_type="mRNA"			
/strain="C57BL/6"			
/db_xref="taxon:10090"			
/clone="IMAGE: 6838773"			

/tissue type="Whole brain"			
/dev stage="1, 5, and 15 days newborn"			
/lab_host="DH10B (T1 phage resistant)"			
/clone_lib="NIH BMAP GH0"			
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. Ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGTGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."			
BASE COUNT	255 a	176 c	189 g
ORIGIN	176 t	3	others
Alignment Scores:			
Pred. No.:	8.7e-71	Length:	799
Score:	1155.00	Matches:	230
Percent Similarity:	97.87%	Conservative:	0
Best Local Similarity:	97.87%	Mismatches:	5
Query Match:	13.94%	Indels:	2
DB:	14	Gaps:	0
US-09-617-099b-1 (1-1590) x CB519497 (1-799)			
Qy	273	ArgLysArgSerProSerValSerArgAspGlnAsnArgArgTyrGluGlnSerGluGlu	292
Db	97	AGAAAAACGAAGTCCATCATCTAGTGTCCAGGATCCAAATCGAAGTACGACGAAAGTGAAGAA	156
Qy	293	ArgGluAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAsp	312
Db	157	AGAGAGGACTACTCACAGTATGTTCTTCAGATGTGTAATGCAAGATCTCTTCGGAT	216
Qy	313	TyrAlaAspArgArgSerGlnArgGluProGlnPheTyrGluGluProGlyHisLeuAsn	332
Db	217	TATGCTGATAGACGATCTCAGCGTGAGCCTCAATTTTATGAAGAACCTGGTCATTAAAT	276
Qy	333	TyrArgAspSerAsnArgGlyHisArgHisSerLysGluTyrIleValAspAspGlu	352
Db	277	TACAGGGATTCTAACAGGAGAGGCCATAGACATTCCTCAAGAGTATATTGTGATGATGA	336
Qy	353	AspValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGlnAlaArg	372
Db	337	GATGTGGAGAGCAGAGATGAATATGAAGACAAAGAGAGAGAGGAAATACCAAGCACGC	396
Qy	373	TyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGln	392
Db	397	TACAGAGTGTATCCAAATCTGGCCGGTATCCGCTAAAGGCCCAACCCCTACCAAGAACAA	456
Qy	393	MetArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSer	412
Db	457	ATGCGCATCCAGCTGAGTGTCCAGGGCAGCATGAGAGAGGACACAGTATGTTTCT	516
Qy	413	LeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspArgProSer	432
Db	517	TTGGCAACGCTGAATAGAGATTCAGAGATTCCTCTCTAAGGATGATAGACCATCA	576
Qy	433	ArgGlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTyrSerMet	452
Db	577	AGGCAAGATCTGATCTGAACGTAGAGCTGCAATGGAACCAACGATCTGATTCATG	636
Qy	453	GluArgThrArgGluAlaGlnGlyGlnSerSerTyrProGlnArgThrSerAsnHisSer	472
Db	637	GAAAGAACTCGAGAGGCTCAGGACANAGTTCTTTATCCAGAGGACCTCANATCATAGT	696
Qy	473	ProProThrProArgArgSerProIleProLeuAspArgProAspMetArgArgAlaAsp	492

Db 697 CCTCCACC-CCTGCGCGGAGCCCTATACCGCTTGATAGACAGACATGAGCGCGCTGAC 755

QY 493 SerLeuArgIysGlnHisHisLeuAspProSerSerAlaValArg 507

Db 756 TCCCTACGGAAACAGCA-CACCTTAGATCCCGAGCTCTGCTGTGAGG 799

RESULT 8
 BB428808
 LOCUS
 DEFINITION
 BB428808 RIKEN full-length enriched, adult male hippocampus Mus musculus cDNA clone CG30023J21.3', similar to AB018294 Homo sapiens mRNA for KIAA0751 protein, mRNA sequence.
 BB428808
 BB428808.2 GI:16423945
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (bases 1 to 686)
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koye, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Sasaki, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 Unpublished
 On Jun 19, 2000 this sequence version replaced gi:9268535.
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
 Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 FEATURES
 Location/Qualifiers
 1..686
 /organism="Mus musculus"
 /mol_type="mRNA"

AUTHORS TITLE JOURNAL COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement:
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: L1AMI4590 row: e column: 18
High quality sequence start: 16
High quality sequence stop: 583.
Location/Qualifiers

FEATURES source

```

1..1004
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6959035"
/dev stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHG_XGC_Eye1"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection (XGC)
) library."
BASE COUNT      320 a      193 c      247 g      243 t      1 others
ORIGIN

```

Alignment Scores:

```

Pred. No.:      3,09e-63      Length:      1004
Score:          1048.50      Matches:      229
Percent Similarity: 86.42%      Conservative: 32
Best Local Similarity: 75.83%      Mismatches:  31
Query Match:      12.66%      Indels:      13
DB:              14          Gaps:         3

```

US-09-617-099B-1 (1-1590) x CD301064 (1-1004)

```

QY      1295 SerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAla 1314
Db      1 AGTGTGAGTGGAGAAAA-TACCAGCTGGAACACAAATGATGTAGTCACTGATACAGCC 59
QY      1315 ValGlyAlaLeuGlyThrSerGlyLysIleArgSerSerIleGlyAlaLysMetVal 1334
Db      60 GTTGGGACGATTGGACCTGTGGAAAGAAAGGAGATCAAGTCTAAGTGCCAAAGTAGTT 119
QY      1335 AlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGly 1354
Db      120 GCTATAGTA-----TACGAGAGAGCAGAGCACATCCCGCTTACCCAAACAGATTCT 173
QY      1355 GlyGlyLysLysLeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGlu 1374
Db      174 GGCACACAGAAATTTAAAGATCTATTCAAGAGAGCAGACAGAAACAGCATGGCCAGCAGAA 233
QY      1375 MetArgAsnTrpMetThrArgGlnAlaSerArgGlnSerThrAspGlySerMetAsnSer 1394
Db      234 ATGAGAAGTCGGATGGTTCGGCACCCGCGAGAGGTCAACAGATGGGAGTATTAAACAGT 293
QY      1395 TyrSerSerGluGlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPhe 1414
Db      294 TACAGCTCAGAGGGCAATTTAAATTTTCCAGGGTGAATTCGGGGTTGATAGCAATTC 353
QY      1415 SerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThr 1434
Db      354 AGCGATTTCTGGATGGACTAGGACCCAGCACAACTTA--GGCAGACAAACGTTAGCAACT 411

```

```

QY      1435 ProAlaMetGlyAspIleGlnValGlyMetMetAspLysLysGlyGlnLeuGluValGlu 1454
Db      412 CCAGCAATGGGGATATTGAGATGGATAGATAAAAGGGGGAGTGTAGAGTGGAA 471
QY      1455 IleIleArgAlaArgGlyLeuValValLysProGlySerIysThrLeuProAlaProTyr 1474
Db      472 GTTATACGAGCAAGAGGTTTAAATACAAAGCCAGGGCTCTAAATCAACTCCAGCTCTTAT 531
QY      1475 ValLysValTyrLeuLeuAspAsnGlyValCysIleAlaLysLysLysThrLysValAla 1494
Db      532 GTAAAGGTATATCTTTTAGAAATGGAGCGTGTGTAGCCAGANAAGACAGAAATGCA 591
QY      1495 ArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGly 1514
Db      592 CGGAAAACCCCTGGACCTCTTTTACCACAGAACTTGGCTTTTGAAGAAAGTCCACATGG 651
QY      1515 ArgValLeuGlnIleIleValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMet 1534
Db      652 AAGGTCTTTGAGTAAATGTCTGGGGAGACTATGGTGAATGGATCGATCACAATCTTTATG 711
QY      1535 GlyValAlaGlnIleLeu-LeuAspGluLeuLeuLeuSerAsnMetValIleGlyTrpPh 1554
Db      712 GGTGTGTGCACAGATCTTTTGGAGAACTCGATCTATCAAGTGTGTGATGGATGGTA 771
QY      1554 e-LysLeuPhePro-ProSerSerLeu-ValAspProThr---SerAlaProLeuThrAr 1572
Db      772 TTAATTTGTTTCGGCCCTTGGTCAATGGTGGATCCACGCTGACACCCCTTACCTCG 831
QY      1572 gArgAla-----SerGlnSerSerLeuGluSerSerThr-GlyProSerTyrSer 1598
Db      832 TAGGGG-CTTCCCAATTCTCTCTTTGAAAGTTTCTACAGGGTCTCTCCCTTTCA 886

RESULT 11
AK043152      1718 bp      mRNA      linear      HTC 05-DEC-2002
LOCUS      Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
DEFINITION      enriched library, clone:A730060M23 product:NIM2 homolog (Rattus
      norvegicus), full insert sequence.
ACCESSION      AK043152
VERSION      AK043152.1 GI:26089509
KEYWORDS      HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1
      Carninci, P. and Hayashizaki, Y.
      High-efficiency full-length cDNA cloning
      Meth. Enzymol. 303, 19-44 (1999)
MEDLINE      99279253
PUBMED      10349636
REFERENCE      2
      Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
      Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
      Normalization and subtraction of cap-trapper-selected cDNAs to
      prepare full-length cDNA libraries for rapid discovery of new genes
      Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL      20493374
MEDLINE      11042159
PUBMED      11042159
REFERENCE      3
      Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
      Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
      Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
      Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
      Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
      Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
      Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
      RIKEN integrated sequence analysis (RISA) system--384-format
      sequencing pipeline with 384 multicapillary sequencer
      Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL      20530913
MEDLINE      11076861
PUBMED      11076861
REFERENCE      4

```

AUTHORS

Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Komno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamana,K., Saito,T., Okazaki,Y., Gojibori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaado,I., Pesole,G., Suzuki,K., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,K., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Balderelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.P., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wyshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

REFERENCE

5 The FANTOM Consortium and the RIKEN Genome Exploration Research

AUTHORS

Group Phase I & II Team.

TITLE

Analysis of the mouse transcriptome based on functional annotation

JOURNAL

of 60,770 full-length cDNAs

REFERENCE

Nature 420, 563-573 (2002)

AUTHORS

6 (bases 1 to 1718)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takada,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE

Direct Submission

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-3222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

FEATURES

Location/Qualifiers
1..1718
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A730060M23"
/db_xref="taxon:10090"
/clone="A730060M23"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="7 days neonate"
498..1423
misc_feature
/note="NIM2 homolog [Rattus norvegicus] (SPTIR[Q9JIR3, evidence: FASTY, 99.3%ID, 100%length, match=923] putative)"

QY 1390 lySerMetAenSerTySerSerGluGlyAsnLeuIlePheProGlyValArgLeuAlas 1410
 Db 828 GAGGACCCACAGTACAGCTCCGAGGCGACGTTATTTTCC--ACACGGCTGGAG 884
 QY 1410 exAspSerGlnPheSerAppPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgG 1430
 Db 885 CTGAAGACCGAGTTCACTGAGTCTGATGGCTGGGACCCAGCCGAGTGTGGGGCGAC 944
 QY 1430 lNThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysGlyG 1450
 Db 945 AGACGCTGGCCACACCCCATGGAGATGTCACATTGTCATCATGAGCCGAGGTGGCC 1004
 QY 1450 lNLeuGluValGluIleAlaArgGlyLeuValValValValValValValValVal 1470
 Db 1005 AGCTGGAGTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1064
 QY 1470 euProAlaProTyValValValValValValValValValValValValValVal 1490
 Db 1065 TCCAGCCACCTATATCAAGCTTACCTGCTGGAGAACGGGCGATGTGTAGCCAGAAA 1124
 QY 1490 yThrLysValAlaAlaArgLysThrLeuGluProLeuTyGlnGlnLeuLeuSerPheGluG 1510
 Db 1125 AGCAAAAGTGGCCAGAAACATATGATCCCTGTACCGACGAGCTCTCTCTTTGATG 1184
 QY 1510 luSerProGlnGlyArgValLeuGlnIleValTyrGlyAspTyGlyArgMetAspH 1530
 Db 1185 AGGGCCCCAAGGCAAGTCTGCGAGTGTCTGCTGGAGAGACTACGGCCGATGGACC 1244
 QY 1530 isLysSerPheMetGlyValAlaGlnIleLeuLeuAspGluLeuGluLeuSerAsnMetY 1550
 Db 1245 ACAAGTCTTCATGGGTATGGCCCGACATCATGCTGGAGAGCTGGACCTGAGTGTCTG 1304
 QY 1550 alledGlyTyrPheLysLeuPheProSerSerLeuValAspProThrSerAlaProL 1570
 Db 1305 TCACCGGCTGTGTAACAACTCTTCCCACTCTCTGTGGCCGACTCCACACTCGGATCC 1364
 QY 1570 euThrArgArgAlaSerGlnSerSerLeuGluSerThrGlyProSerTySer 1588
 Db 1365 TCACCGGCGCTCTCTCAGTCTCTTGGAGTGCACACGCCCCCTCTGTCTCT 1420

RESULT 12
 BU741736/c
 LOCUS BU741736 697 bp mRNA linear EST 10-OCT-2002
 DEFINITION UI-E-EOI-aiz-h-20-0-UI.s1 UI-E-EOI Homo sapiens cDNA clone
 UI-E-EOI-aiz-h-20-0-UI 3', mRNA sequence.
 BU741736
 ACCESSION BU741736.1 GI:23687327
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 697)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 FORWARD

FEATURES
 Location/Qualifiers
 1..697
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-EOI-aiz-h-20-0-UI"
 /tissue_type="fetal eye"
 /dev_stage="fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-EOI"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-EOI is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
 TAG_LIB=UI-E-EOI
 TAG_TISSUE=human fetal eye
 TAG_SEQ=CCGTATACC"
 BASE COUNT 164 a 161 c 159 g 212 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1,28e-62 Length: 697
 Score: 1037.00 Matches: 210
 Percent Similarity: 98.60% Conservative: 1
 Best Local Similarity: 98.13% Mismatches: 1
 Query Match: 12.52% Indels: 2
 DB: 13 Gaps: 0

US-09-617-099b-1 (1-1590) x BU741736 (1-697)

QY 1379 MetThrArgGlnAlaSerArgGluSerThrAspGlySerMetAenSerTySerSer-Gl 1398
 Db 696 ATGACTCGACAGCAAGCCGAGAGTCTACAGATGGTAGTGAACAGCTACAGCTCAGGA 637
 QY 1398 uGlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLe 1418
 Db 636 AGGAATCTGATTTTCCCTGGTGTCTGCTGGCTCTGATAGCCAGTTCACTGATTTCTCT 577
 QY 1418 uAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGl 1438
 Db 576 GGTATGGCTTGGCCCTGCTCAGCTAGTGGGACCCAGACTCTGGCAACACCTGCAATGGG 517
 QY 1438 yAspIleGlnValGlyMetMetAspLys-LysGlyGlnLeuGluValGluIleLeArgA 1458
 Db 516 TGACATTTCAGTGGATGATGACAAANAAGGACAGCTGGAGGTAGAAATCATTCGCGG 457
 QY 1458 laArgGlyLeuValValValValValValValValValValValValValValValVal 1478
 Db 456 CCGTGGCTTGTGTGTAAACACAGGTTCACAGACACTGCCAGCACCTGTATGTAAAGTGT 397
 QY 1478 yrLeuLeuAspAsnGlyValCysIleAlaLysLysLysLysLysLysLysLysLysLys 1498
 Db 396 ATCTATTAGATAACGGAGTCTGCATGCCAAAAGAAAACAAAAGTGGCAAGAAAACGC 337
 QY 1498 euGluProLeuTyGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuG 1518
 Db 336 TGAACCCCTTTTACCAGCAGCTATTATCTTTTCGAAGAGAGTCCACAGGAGAAATTTTAC 277
 QY 1518 lNleIleValTyrGlyAspTyGlyArgMetAspHsLysSerPheMetGlyValAlaG 1538

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs@mail.nih.gov
 Tissue Procurement: the Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lari, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 81 Row: p Column: 13
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, GenomScan gene prediction
 This clone has the following problem: retained intron.

FEATURES

source
 Location/Qualifiers
 1..3080
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:5400144"
 /tissue_type="Eye, retina, mouse strain C57BL/6"
 /clone_lib="NIH MGC_94"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 BASE COUNT 915 a 648 c 632 g 885 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,86e-60 Length: 3080
 Score: 1006.50 Matches: 195
 Percent Similarity: 90.34% Conservative: 20
 Best Local Similarity: 81.93% Mismatches: 22
 Query Match: 12.15% Indels: 1
 DB: 11 Gaps: 1
 US-09-617-099B-1 (1-1590) x BC037619 (1-3080)

Qy 1353 GluGlyGlyLysLysLeuArgSerThrValGlnArgSerThrGluThrGlyLeuAla 1372
 Db 123 GAGTCCGGCCACAGAAAGTTAAAAGACCACTTACAGAGGAGTACGGAACGGGAATGGCG 182
 Qy 1373 ValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerThrAspGlySerMet 1392
 Db 183 GCTGAATGGGAAG---ATGGTGAGACAGCCCGGGAGTCCAGGATGGCGAGTC 239
 Qy 1393 AsnSerTyrSerSerGluGlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSer 1412
 Db 240 AACAGTTATAGCTCAGAGGAAGAACTTAATTTCTCGAGTTCGCGGGGACCTGACAGT 299
 Qy 1413 GlnPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeu 1432
 Db 300 CAGTTTCAGTGATTTCCITGATGGATTGGGGCCAGCCAGCTTGTGCGCGCCCAACCCCTC 359
 Qy 1433 AlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLysGlyGlnLeuGlu 1452
 Db 360 GCCACCCAGCCATGGGCGATATCCAAATCGGATGGAGGATAAGAGGGTCAGTTGGAG 419
 Qy 1453 ValGluIleIleArgAlaArgGlyLeuValValLysProGlySerLysThrLeuProAla 1472
 Db 420 GTTGAGGTTATCAGAGCCCGGAGCCTTACACAAAACCTGGCTCCAAATCTACACCTGCT 479

Qy 1473 ProTyrValLysValTyrLeuLeuAspAsnGlyValCysIleAlaLysLysLysThrLys 1492
 Db 480 CCTATGTGAAGTTTATCTTTTGGAAATGGAGCTGTATGGCCAAAAGAGACAGA 539
 Qy 1493 ValAlaArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerPro 1512
 Db 540 ATTGCACGGAAACTCTCGATCCCTTGATCAGCAGTCCCTGTTTGTGATGAAAGTCCA 599
 Qy 1513 GlnGlyArgValLeuGlnIleIleValTrpGlyAspTyrGlyArgMetAspLysSer 1532
 Db 600 CAGGGTAAAGTTCTTCAGGTGATTGTCTGGGTGACTATGGAAGAAATGACCAAAATGC 659
 Qy 1533 PheMetGlyValAlaGlnIleLeuLeuAspGluLeuGluLeuSerAsnMetValIleGly 1552
 Db 660 TTTATGGGTGTGGCTCAATCTTGTGGAAAGAACTTGATCTGTCCAGCATGTTGATGGA 719
 Qy 1553 TrpPheLysLeuPheProSerSerLeuValAspProThrSerAlaProLeuThrArg 1572
 Db 720 TGTATATAAATTGTTCCCTCATCTCTCTGTTGGATCCACACTCACTCCCTGACCCGC 779
 Qy 1573 ArgAlaSerGlnSerSerLeuGluSerThrGlyProSerTyrSerArgSer 1590
 Db 780 CGGGCTTCCCAATCATCACTGGAAGTTCTGTGGGCTCCCTGCAATCCGGTCA 833
 RESULT 15
 LOCUS CA346497 663 bp mRNA linear EST 05-NOV-2002
 DEFINITION 677393 NCCOWA 1RT Oncorhynchus mykiss cDNA clone lrt101d11_B_06
 5', mRNA sequence.
 ACCESSION CA346497
 VERSION CA346497.1 GI:24591668
 KEYWORDS EST.
 SOURCE Oncorhynchus mykiss (rainbow trout)
 ORGANISM Oncorhynchus mykiss
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 REFERENCE 1 (bases 1 to 663)
 AUTHORS Rexroad, C.E. and Keele, J.W.
 TITLE Sequence analysis of a rainbow trout normalized cDNA library
 JOURNAL Unpublished
 COMMENT Contact: Rexroad CE
 USDA, ARS, National Center for Cool and Cold Water Aquaculture
 11876 Leetown Road, Kearneysville, WV 25430, USA
 Tel: 304 724 8340 x2129
 Fax: 304 725 0351
 Email: crexroad@nccowa.ars.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified by
 cross_match v0.990329.
 Seq primer: ATTAGTGACACTATAG.
 Location/Qualifiers
 source
 1..663
 /organism="Oncorhynchus mykiss"
 /mol_type="mRNA"
 /db_xref="taxon:8022"
 /clone="lrt101d11_B_06"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="NCCOWA 1RT"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from brain, gill, liver,
 spleen, muscle, and kidney."
 BASE COUNT 170 a 136 c 173 g 124 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.41e-60 Length: 663
 Score: 999.00 Matches: 192
 Percent Similarity: 91.28% Conservative: 7
 Best Local Similarity: 88.07% Mismatches: 3
 Query Match: 12.06% Indels: 16

DB:	14	Gaps:	1
US-09-617-099B-1 (1-1590) x CA346497 (1-663)			
Qy	753	ProGluProGlnValGluLeuValSerArgProIleGlyAspIleProArgIlePro	772
Db	3	CCAGAACCCGAGTGGAACTGGTGGTCTCTAGAGCCCATAGGAGATATTCCTCGATACCA	62
Qy	773	AspSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGlnLysMetAsp	792
Db	63	GACAGCACACATGCACAACATCGAATCAAGTTCGAGTTCCCTTGAGTCTCAGAGATGGAT	122
Qy	793	ArgProSerIleSerValThrSerProMetSerProGlyMetLeuArgAspValProGln	812
Db	123	CGCCCTCTATCTCTGTGTCACGTCTCCCATGAGCCCGGGATGCTGCGGAGCGCCCCCAG	182
Qy	813	PheLeuSerGlyGlnLeuSer-----	819
Db	183	TACCTGTCCAGACAGCTCTCAAGCCAAAGCCTTAGTAGAACAAGCGCCTTTGTCCCT	242
Qy	820	-----IleLysLeuTrrPheAspLysValGlyHisGlnLeuIleValThrIleLeu	836
Db	243	AGGGTCCAGGTCAAACTGTGGTATGACAAAGTGGGCCATCAGCTAATAGTACCATCCTG	302
Qy	837	GlyAlaLysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIle	856
Db	303	GGAGCCAAAGGACCTGCCTTCTAGGGAAGATGGCCGCCCGGAAACCTTAGCTCAAAATC	362
Qy	857	TyrPheLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThr	876
Db	363	TACTTCTCCTTGACAGAGCGGACAAAGCAAGAGAGAGAGACAAACCGTAAGAAATCC	422
Qy	877	LeuGluProLysTrrAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArg	896
Db	423	TTAGAGCCCAATGGAACAGACCTTATGTATTCTCCTGTGCACCCGGGGAGTTCCGT	482
Qy	897	GluArgMetLeuGluIleThrLeuTrrAspGlnAlaArgValArgGluGluSerGlu	916
Db	483	GAGCGCATGTTGGAGATCACCTGTGGGACCCAGGCCAGGGTCAGGGAGAGAAAGCGAG	542
Qy	917	PheLeuGlyGluIleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrr	936
Db	543	TTCTCTGGGAGAGATTAAGTAGAGCTGGAGACAGCCCTGCTGGATGATGAGCTCCTGG	602
Qy	937	TyrLysLeuGlnThrHisAspValSerSerLeuProLeuProArgProSerPro	954
Db	603	TACAGCTGCAGACCCACAGCGTGTCTCCATGCCCTCCCGCCCTCGGCC	656

Search completed: November 16, 2003, 15:02:58
Job time : 6846 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 09:19:06 ; Search time 1020 Seconds
(without alignments)
5094.554 Million cell updates/sec

Title: US-09-617-099B-1

Perfect score: 8285

Sequence: 1 MSAPLGRGRPAFTPAASQP.....TTRASQSSLESSTGTPSYRS 1590

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 433922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09617099/runat_14112003_191547_23738/app_query.fasta_1.1735
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpp -MINMATCH=0.1
-LOOPCI=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09617099@cgn 1 1 708 @runat_14112003_191547_23738
-NCPU=6 -ICPU=3 -NO MAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	8285	100.0	4980 14	US-10-180-326-2 Sequence 2, Appli

2	2936.5	35.4	6638	12	US-10-220-891-40	Sequence 40, Appl
3	1042	12.6	7325	12	US-10-240-965-72	Sequence 72, Appl
4	374	4.5	9027	12	US-09-930-213-304	Sequence 304, App
5	368.5	4.4	24387	11	US-09-764-891-9844	Sequence 9844, Ap
6	332	3.8	9025608	14	US-10-156-761-13	Sequence 1, Appli
7	300.5	3.6	9169	12	US-09-814-353-20083	Sequence 20083, A
8	295.5	3.6	9025608	14	US-10-156-761-13	Sequence 1, Appli
9	295	3.6	9161	12	US-10-133-937-37	Sequence 37, Appl
10	295	3.6	11950	12	US-10-252-157-218	Sequence 218, App
11	288.5	3.5	52101	12	US-10-132-134-11	Sequence 1, Appli
12	284.5	3.4	4215	14	US-10-177-293-448	Sequence 448, App
13	280	3.4	11295	12	US-09-814-353-19153	Sequence 19153, A
14	279	3.4	12227	14	US-10-177-293-442	Sequence 422, App
15	270.5	3.3	5418	14	US-10-156-761-5959	Sequence 5959, Ap
16	270.5	3.3	7997	9	US-09-864-761-17726	Sequence 17726, A
17	268	3.2	9579	12	US-10-132-134-11	Sequence 11, Appl
18	265.5	3.2	10383	12	US-09-960-253-181	Sequence 181, App
19	265.5	3.2	10383	12	US-10-007-926A-56	Sequence 56, Appl
20	264.5	3.2	15872	10	US-09-861-289-1	Sequence 1, Appli
21	264.5	3.2	15872	10	US-09-860-846-1	Sequence 1, Appli
22	264.5	3.2	15872	10	US-09-988-384B-1	Sequence 1, Appli
23	264.5	3.2	15872	11	US-09-836-821-1	Sequence 1, Appli
24	264.5	3.2	15872	12	US-10-271-883-44	Sequence 44, Appl
25	264	3.2	77536	12	US-09-940-316B-1	Sequence 1, Appli
26	260.5	3.1	68750	13	US-10-014-717-1	Sequence 1, Appli
27	260	3.1	10174	14	US-10-171-311-82	Sequence 82, Appl
28	260	3.1	42999	10	US-09-799-462A-17	Sequence 17, Appl
29	260	3.1	42999	11	US-09-836-911A-17	Sequence 17, Appl
30	260	3.1	42999	12	US-09-738-630-73	Sequence 73, Appl
31	260	3.1	42999	13	US-10-125-767-17	Sequence 17, Appl
32	260	3.1	42999	14	US-10-151-081-17	Sequence 17, Appl
33	260	3.1	42999	14	US-10-287-313-17	Sequence 17, Appl
34	260	3.1	42999	14	US-10-219-694-17	Sequence 17, Appl
35	260	3.1	88421	10	US-09-976-059-1	Sequence 1, Appli
36	259	3.1	3651	12	US-10-311-406-3	Sequence 3, Appli
37	257	3.1	38734	11	US-09-373-658-30	Sequence 30, Appl
38	256	3.1	14835	12	US-10-240-965-113	Sequence 113, App
39	255.5	3.1	6457	10	US-09-880-107-3389	Sequence 3389, Ap
40	255.5	3.1	7497	10	US-09-960-253-175	Sequence 175, App
41	255.5	3.1	10056	14	US-10-156-761-411	Sequence 411, App
42	255.5	3.1	100000	14	US-10-156-761-15103	Sequence 15103, A
43	254.5	3.1	7058	12	US-10-287-218-38	Sequence 38, Appl
44	254.5	3.1	7792	13	US-10-044-090-359	Sequence 359, App
45	254	3.1	4974	14	US-10-156-761-2487	Sequence 2487, Ap

ALIGNMENTS

RESULT 1
US-10-180-326-2
; Sequence 2, Application US/10180326
; Publication No. US20030049661A1
; GENERAL INFORMATION:
; APPLICANT: Seino, Susumu
; APPLICANT: Shibusaki, Tadao
; APPLICANT: Ozaki, No. US20030049661A1uaki
; TITLE OF INVENTION: Protein Rim2
; FILE REFERENCE: P21573
; CURRENT APPLICATION NUMBER: US/10/180,326
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: JP 288372/99
; PRIOR FILING DATE: 1999-10-08<160> 5
; NUMBER OF SEQ ID NOS:
; SOFTWARE: Patent version 3.0
; SEQ ID NO 2
; LENGTH: 4980
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-180-326-2

Alignment Scores: 0 Length: 4980
Pred. No.: 8285.00 Matches: 1590
Score:

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-617-099B-1 (1-1590) x US-10-180-326-2 (1-4980)

Qy 1 MetSerAlaProLeuGluProArgGlyArgProAlaProThrProAlaSerGlnPro 20
Db 32 ATGTGGCTCGCTCGGGCCGCGGGCCGCGCTCCACCCCGCGGCTCTCAACCT 91
Qy 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysIlelle 40
Db 92 CCTCGCAGCCGAGATGCCGACCTCAGCCACCTCAGAGAGAGAGAGAAATCATC 151
Qy 41 LeuAlaValMetAspArgGlnLysGluGluGluLysGluGlnSerValLeuLysIle 60
Db 152 CTGGCTGTCATGATCGTCAGAGAAAGAGAGAGAGAGAGAGATCCGCTCAAGATC 211
Qy 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
Db 212 AAAGAAGAACACAAAGACACACCGACACAGTGGTTTCCTTTAGTGGGATCACTGAAC 271
Qy 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
Db 272 GTAATAACGTTCTCGAGCCCGCAGCAAAACCAACCAATGAGAGAGAGAGAGAGAG 331
Qy 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120
Db 332 CTGCACCAACATTTGAAATGATTAAGGAGCAAGTCAAGAGATGGGAGAGAAATCGCAG 391
Qy 121 GlnGlnGlnGluLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
Db 392 CAGCAGCAGACAGAGAGGTGATCCCGACCTGTGGCATCTGCCACAGACAAATTT 451
Qy 141 AlaAspGlyCysGlyHisLysCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
Db 452 GCAGATGGATGGCGCCATTAATTTGTTCTTATGCCAAACCAAGTTCTGTGCTCGATGGA 511
Qy 161 GlyArgValSerLeuArgSerAsnLysValMetTyrValCysAsnLeuCysArgLysGln 180
Db 512 GGTGAGTGCTCTTAGCTCAACCAAGTTATGTGGTGTTGTAATTTGGCCGAGAAACAA 571
Qy 181 GlnGluLeuLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
Db 572 CAAGAAATCCTCAATAATCAGAGCATGGTTTATATAAGTGGGTCTAACACACTGCAG 631
Qy 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGluLysLys 220
Db 632 CAACCTGATCAAAAGGTTCCTCGAGGGCTTCGAAATGAGAGAGCCCTCAGAGAGAGAAA 691
Qy 221 AlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAla 240
Db 692 GCAAAACTACAGCAGCAGCCCGATTTCCAAAGAGAGCCCGAGGTGACTTATCAGTACCTGCA 751
Qy 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
Db 752 GTTGAAAGGGCGAGCTCATGGGCTCACAAGACAGGATACTATTAAAAATGGATCAGGA 811
Qy 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
Db 812 GTGAAGCACAGATGCCAGTGCATGCCCTTCAGACAGAGAAACGAAAGTCCATCAGGTGCC 871
Qy 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrVal 300
Db 872 AGGGATCAAAATCGAGATACGAGCAAGTGAAGAGAGAGAGAGACTACTACAGATATGTT 931
Qy 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSerGlnArg 320
Db 932 CCTTCAGATGGTACAATGCCAAGATCTCTTCGGATTTATGCTGATAGACGATCTCAGCGT 991
Qy 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340

Db 992 GAGCCTCAATTTTATGAAGAACCTGGTCAATTTAAATTTACAGGGATTCTTAACAGGAGAGC 1051
Qy 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
Db 1052 CATAGACATTCCAAAGAGATATTTGTGGATGATGAGATGTGGAGACAGAGATGAATAT 1111
Qy 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
Db 1112 GAAGACAAAGGAGAGAGAGAGATACAGGCAGCTACAGAGTGTATCAAAATCTCGCC 1171
Qy 381 ArgTyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValSer 400
Db 1172 CGGTATCCCGTAAAGCCACCACTACGAAGAACAAATGGCATCCACGCTGAGGTGCC 1231
Qy 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
Db 1232 AGGCACGACATGAGAGAGAGCAAGTGTGTTCTTTGGCAACGCTGAACATAGAAGAT 1291
Qy 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
Db 1292 TCCAGGATTTCTCTGCTAAGGATGATAGACCATCAAGGCAAGATCTGTATCTGAACGT 1351
Qy 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
Db 1352 AGAGCTGCAATGGAAACCAACGATCGTATTCAATGGAAAGAACTCGAGAGGCTCAGGGA 1411
Qy 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
Db 1412 CAAGTTCCTTATCCACAAAGACCTCAATCATAGTCTCTCCACCCCTCGCGGAGCCCT 1471
Qy 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
Db 1472 ATACCGTTGATAGACAGACATGAGGCGCGCTGACTCCCTACGGAAACAGACACCTTA 1531
Qy 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
Db 1532 GATCCAGCTCTGCTGTGAGGAAACCGAAGCGAGAAAAATCGAAACCATGTGTAAGGAAT 1591
Qy 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProArgProHisLys 540
Db 1592 GATTTCTTTGAGTTCCAGACCGTCCGAGTCAGTGAGCGCCGCCCAACCAAGGCTCATAAA 1651
Qy 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAla 560
Db 1652 TCCAAAGAAAGAGGTAAATTCGCCCGCTTCTACTGACAGCTCGAGAGAGAGTGCGCA 1711
Qy 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
Db 1712 TCCACACCTGAGTATACAAAGCTGTGATGAGAGCTCGAAAGCGAGAGTGTGAGTGAG 1771
Qy 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
Db 1772 AAAGGGGACAGTCAAAAGGGAAAAAGAAAACTAGTGAGCAGGAGGTTTGTGCGATCT 1831
Qy 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
Db 1832 AACACAGGTCTGAGAGACAAAGAAAGAGTACTATGTTGGCCACTCTTTTGGAGAGAG 1891
Qy 621 AspLeuGluTyrSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640
Db 1892 GATTTGGAATGCTCTGAGCCTCAGATTAAGGACTCTGGGTAGATACCTGTAGTAGCACA 1951
Qy 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
Db 1952 ACCCTTAACGAGGAGCATAGCCATAGTAGACCCCTGTGACCTGGCAGCCATCCAAA 2011
Qy 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 2012 GATGAGATCGCCTCAATTTGGTGTATTTTATTAATAGCGTTTAAAGATGGGAGTGTA 2071
Qy 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
Db 2072 CTTCCAGACTCAGGAGCAATGCTGGGCTTAAGGTTTAGGAGAGAAAGATGACTGAATCA 2131

QY 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
DB 2132 GGTGGACTTTGTGCAATTATATACCAAGTAAAGAGAGATTAGCTGATCTGTAGGA 2191
QY 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr 740
DB 2192 CATCTTAGACCGAGTGATGAAGTCTTGGAAATGGAGGCTATTTCGAAGGAGCCACA 2251
QY 741 PheGluGluValTyrAsnIleLeuLeuSerLysProGluProGlnValGluLeuVal 760
DB 2252 TTTGAGAGATTATACAACTATTCTAGAAATCCAAACCTCAACCAAGTTGAGCTTGT 2311
QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
DB 2312 GTTTCAGGCGCAATTGGAGATATCTTAGAATACCTGATAGCAGCGATGCACAACTGGAA 2371
QY 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
DB 2372 TCCAGTTCTAGTCAATTTGAATCTCAAAAAATGGACCGTCTCTTATATCCGTTACCTCA 2431
QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
DB 2432 CCCATGATCTCTGCGATGCTGAGGATGTCCCGCAGTTCTTATCTGACAGCTTTCAATA 2491
QY 821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
DB 2492 AAATCTAGTTTGAAGGTTGTGCACAGTTGATGATTAACAATTTTGGAGCAAGGAT 2551
QY 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
DB 2552 CTCCTCCAGGGAAGATGGAGGCCAAGAAATCCTTATGTAAATTTACTTCTCTCCA 2611
QY 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
DB 2612 GATAGAAAGTATATAAAATAAGAGAGAACAAAAACAGTCAAGAAAACTTTGGAAACCCAAA 2671
QY 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
DB 2672 TGGAAACAGACTTTTCATTTATCTCTGTCCACGAGAGAAATTCCTGTGAACGAATGCTG 2731
QY 901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluLeuSerGluPheLeuGlyGlu 920
DB 2732 GAAATTAACCTTTGGGATCAAGCTAGAGTTCCAGAGAGAGAGAGCGGAATCTTAGAGAG 2791
QY 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisIleTyrLysLeuGln 940
DB 2792 ATTTTAATTGAATTTGAAACACAGCTTTGCTAGATGATGAGCGCACCTGGTATAAGCTGCAG 2851
QY 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
DB 2852 ACCCATGATGTCCTCATTTGCCACTCCTCGCCCTTCCCATATCTGCCCGGAGGAG 2911
QY 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
DB 2912 CTCATGGAGAGACCCCAACCGCAGGCTGCAAGAGTTCGAGAGAAATAAGTCAGTGAA 2971
QY 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
DB 2972 GTGTCTGACTACGACTGCGAGATGGCGTGGAGTAGTGTCCAGATTATCGACACAAATGGC 3031
QY 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
DB 3032 CGCATCTTCAAGCTCCACGTTGTGCGTGCACAAACAGTCATGTCATCAAAATCATTGC 3091
QY 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
DB 3092 TCACATCAGGGTCTCTCTCATCGAGTAGATTATAGGAAGACACAGGTCAATGTCGCT 3151
QY 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
DB 3152 AGTGCCCTCTCTCTCAAGGAATGTGAAACAGGGGCACCGGGGACACGCTGCTACTGGC 3211

QY 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
DB 3212 CATTTACACACATATTAGCCGAATGGATAGACACCGGTGTGATGATGCCACTACTCTTCA 3271
QY 1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
DB 3272 GATAGAGACAGGGAATTTGTGAAGCAGCAGATAGACAGCCATATCACAGATCCAGATCAACA 3331
QY 1101 GluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProAsp 1120
DB 3332 GAAACAAGGCTCTCTCTAGAGCGGACACACCCGCTCCAGATCCTCTGAACGTCCTGAT 3391
QY 1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140
DB 3392 ACAACCTCATGAGGTGATGCTTCAATTAAGTGAAGATCTGCCCTCTCTCACCT 3451
QY 1141 AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrPro 1160
DB 3452 GCCTTATCGAGGTCTCACCTCGTACCGGTCTGTCCAGACAGCCCATCAAGTACTCCG 3511
QY 1161 GlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArg 1180
DB 3512 GGAACAGGACGAAGGGCGGACAGCTTCCACAGCTTCCACAAAGGGAACATTTGGAGAGA 3571
QY 1181 SerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnVal 1200
DB 3572 AGTGCTATGGATATAGAGAGAGAAATCGCCAAATGAACCTTAACAAATACAAACAGGTA 3631
QY 1201 AlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpAsp 1220
DB 3632 GCCGGATCAGACCCAGACTGGAGCAAGATTACCAITTCGAAGTATCGCTCAGGATGGAT 3691
QY 1221 ProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSerAspVal 1240
DB 3692 CCACATAGAGGGCGAGATACTGTTTCCATTAATCTCCGACAGTATGATGAAGTGATGA 3751
QY 1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerVal 1260
DB 3752 TCTCGGTTTCAAGGACTAGTAGTGTCTTCGTTTCAGCAGCAGCAAGCTACATGTCGCTC 3811
QY 1261 GlnSerGluArgProArgGlyAsnArgLysValSerValPheThrSerLysMetGlnAsn 1280
DB 3812 CAATCAGAGCGGCGGAGAGAAACAGAGAAATCAGTGTCTTTACATCCAAAATGCAAAAC 3871
QY 1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerLysSerGlyAspMet 1300
DB 3872 AGACAGATGGCGCTGTGCGGGAAGAACTTCACAAAGACACAGCATCAGTGAGAGACATG 3931
QY 1301 CysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThr 1320
DB 3932 TGCTCACTGGAGAGAAATGACCGGACGCTCCAGCACTGCGGCGCCCTGGGTACC 3991
QY 1321 SerGlyLysLysArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSer 1340
DB 3992 AGTGGCAAGAACCGCGATCTAGCATTTGGGGCCAAAATGGTAGCTATTGTTGGTCTCTCA 4051
QY 1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyLysLysLeuArg 1360
DB 4052 CGGAAAAGTCCGAGTGCCTCTCAACTCAGCAAAACCGAAGGAGGAGTAAAAAGCTACGG 4111
QY 1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
DB 4112 AGCACTGTTCAGAGAGACACGAGACCGGCTAGCAGTGAGATGAGGAAGTATGATGACC 4171
QY 1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsn 1400
DB 4172 CGCCAGGCCAGCGGGAATCCACAGATGGCAGCATGAACAGCTATAGTTCGGAAGGAAT 4231
QY 1401 LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly 1420
DB 4232 CTGATCTCTCCCTGGGTCCTCGCCTCTGACAGCCAGTTCAGTGTGATTTCTCTGGATGGC 4291
QY 1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440

Db	4292	CTGGGCGCTGCTCAGCTAGTGGGACGCCAGACCTCGCTACTCTCTCAATGGGTGCATT	4351
Qy	1441	GlnValGlyMetMetAspLysGlyGlnLeuGluValGlnIleIleArgAlaArgGly	1460
Db	4352	CAGTGGGAATGATGGATAAAGGGACAGCTGGAGGTAGAAATCATCCGGCGCGCGC	4411
Qy	1461	LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu	1480
Db	4412	CTTGTGTTAAACCAGGTTCCAAAGACACTGCAGCAGCCGTATGTCAAGGTGTATCTGTTA	4471
Qy	1481	AspAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeuGluPro	1500
Db	4472	GACAACGGAGTCTGCATAGCCAAAGAAACCAAGGTGGCGAGAAAGACCTCGAGGCC	4531
Qy	1501	LeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleIle	1520
Db	4532	CTGTACCAGCAGCTCTTGCTCTCGAGGAGAGCCCCAGGGGAGGGTGTACAGATCATT	4591
Qy	1521	ValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu	1540
Db	4592	GTCTGGGGAGATTATGGTTCGTATGGATCACAAATCCTTTATGGGAGTGGCCACAGATCTC	4651
Qy	1541	LeuAspGluLeuGluLeuSerAsnMetValIleGlyTrpPheLysLeuPheProProSer	1560
Db	4652	TTAGATGAACCTGGAACTATCCAAATGGTGAATGGATGGTTCAAACTCTTCCTCCCTTCC	4711
Qy	1561	SerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGlu	1580
Db	4712	TCCCTAGTAGATCCAACTCGGCACCTCTGACCAAGAGAGCTTCCCAATCGTCTCTGAA	4771
Qy	1581	SerSerThrGlyProSerTyrSerArgSer	1590
Db	4772	AGTCTACCGGACCTTCTTACTCTCGTTCA	4801

RESULT 2

```

US-10-220-891-40
; Sequence 40, Application US/10220891
; Publication NO. US20030207286A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWARA, AKIRA
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES HAVING CHARACTERISTICS OF ENHANCED
; TITLE OF INVENTION: EXPRESSION IN HUMAN NEUROBLASTOMA WITH FAVORABLE PROGNOSIS
; TITLE OF INVENTION: BASED ON COMPARISON BETWEEN HUMAN NEUROBLASTOMA WITH FAVORABLE
; TITLE OF INVENTION: PROGNOSIS AND HUMAN NEUROBLASTOMA WITH UNFAVORABLE PROGNOSIS
; FILE REFERENCE: 7388-73435
; CURRENT APPLICATION NUMBER: US/10/220,891
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: JP 2000/140387
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: JP 2000/159195
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 6638
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-220-891-40

```

Qy	21	ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysIleLeu	40
Db	369	CCCCCATGCAAGAGCTGCGCGACCTTCCGCAAGAGAGAGGAACATATATC	428
Qy	41	LeuAlaValMetAspArgGlnLysLysGluGluGluLysGlnSerValLeuLysIle	60
Db	429	ATGGCAGTGATGACCGCGAAGAGGAGGAGGAGAAAGAGAAGCAACCATGTCTCAAGTGT	488
Qy	61	LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu	80
Db	489	GTTGTCCAGGACATGGCGAAGCTGTCTGCC-----	518
Qy	81	ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGlnLysGluProGlnThrLys	100
Db	519	-----TCCAAACACCAAGAAATGCTGAAAAACCGCCGCCCAACCACTTCCACGAGA	569
Qy	101	LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGluGluGluSerGln	120
Db	570	TTGCATCAACAGATTGAAGACTATAAGAAACAAGTGAAGAAATATAGGGGAAGCGCGG	629
Qy	121	GlnGlnGln---GluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLys	139
Db	630	CGTTACACGGCGAGACAAGACGATGCTCCGACTTGTGAATCTGTCTATAAAACAAG	689
Qy	140	PheAlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCys	159
Db	690	TTTGCTGATGGTGGCGTCATCTCTGCTCTCTATTGTCCGACTAAGTTCTGTGGCGCGTGC	749
Qy	160	GlyGlyArgValSerLeuArgSerAsnLys-----ValMetTrpValCysAsn	175
Db	750	GGAGCGCGCGTGTCTTACCGGTCAACCAACGAGACAAAGTGGTTATGTGGGTATGCAAT	809
Qy	176	LeuCysArgLysGlnGlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsnSerGly	195
Db	810	TTATGTCGAAGCAACAGAAATCTTTAACTTCTGGGCATGGTTCTTTTGAAGTGGC	869
Qy	196	SerAsnThrLeuGlnGlnProAspGlnLysValProArgGlyLeuArgAsnGlu-----	213
Db	870	-----CCTCAGCAGCAAGTCAGGATGCAACCCCTGAGTGATACAGCT	911
Qy	214	-----GluAlaProGlnGluLysIleLysLeuHisGlnGlnProGln	228
Db	912	ACAGGTGTGGCTCTCTGAGTACCAAGAGAAAGAACGACACTCCAGAGCGCATCGCGG	971
Qy	229	PheGln-----GlyAlaProGlyAspLeuSerValProAlaVal	241
Db	972	TCTCAGACACCCCTGAGCAGCAGCTGCTCTCTCCAGGATGCTGTCTCTCCACGCGCA	1031
Qy	242	GluLysGlyArgAlaHisGlyLeu-----ThrArgGlnAspThrIleLysAsnGlySer	259
Db	1032	CCACCAGACAGGACAAGGGCTGAGCCCTCGCAGCAAGCCTTGGGCGCTGAACAGAAG	1091
Qy	260	GlyValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerVal	279
Db	1092	CAGGCTTCATCCAGGCTAGAAGTGAACCTCTAGAGAGAGAAAGAGACCCCAAGGCTT	1151
Qy	280	SerArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyr	299
Db	1152	TCCGAGCAGATGGCAAGAGAGCCCTTAAGACCGAGCGGAACCC-----	1196
Qy	300	ValProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGln	319
Db	1197	GTGCCAAGACCTCAGCCACCGCTGGAGGGGGCCGTCGAGAACCGGAGCGCAAGAA	1256
Qy	320	ArgGluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArg	339
Db	1257	AGG-----CGGAAAGCCGGAAGGCTT	1277
Qy	340	GlyHisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGlu	359
Db	1278	GAGAAAGGGCGCATCACAGATTAC-----CCAGACACCGCGGAAAAACGGGATGAG	1328

Db 3291 GATCATAGAACACAGAGATCTGATAGTACGTATTTATTCAGAACACAGACAGTCTCTT 3350
Qy 1088 AlalaAspArgGlnProTyrHisArgSerArgSerThrGlu 1101
Db 3351 ATGCTGCCACAGACA-----AAACGAGGACGAAGTGCAGAA 3386

RESULT 3
US-10-240-965-72
; Sequence 72, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, Dov
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAWN, Richard M.
; APPLICANT: SEILHAMER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITA, Thomas
; APPLICANT: TAI, Julie
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 72
; LENGTH: 7325
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030165924A1 977757.3
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 4103, 4411, 4430
; OTHER INFORMATION: a, t, c, g, or other
US-10-240-965-72

Alignment Scores:
Pred. No.: 7,94e-69 Length: 7325
Score: 1042.00 Matches: 245
Percent Similarity: 62.65% Conservative: 62
Best Local Similarity: 50.00% Mismatches: 113
Query Match: 12.58% Indels: 70
DB: 12 Gaps: 13

US-09-617-099B-1 (1-1590) x US-10-240-965-72 (1-7325)

Qy 1125 ArgSerMetPro-----SerLeuMetThrGlyArgSerAlaProSerProAla 1141
Db 66 CGGGCGATGCCCCGTGAGCTCCCTC-----GCCGCCCTCCCGCCCGCG 113

Qy 1142 ---LeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerThrPro 1160
Db 114 TGCTATCCACTCGGATCGCGCCAGCCTGGG-----GCCGGCGCGCCCTACTGCCG 167

Qy 1161 GlyThr-----GlyArgGlyArgGlnLeuProGlnLeu-----ProProLysGly 1176
Db 168 GGTTCGGGGGGGGGTCCTCGGGGACACACTTCGCCCGCCCTTCGGAGCGCCCTCGG--- 224

Qy 1177 ThrLeuGluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLys 1196
Db 225 CCTGTGGAGGCCCTCCCTG-----245

Qy 1197 TyrLysGlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArg 1216
Db 246 -----TCTGGACCCCGGCCCCACCTCCGACCCCTTTTATCATCATCGC 287

Qy 1217 SerGlyTrpAspProHisArgGlyAlaAspThrValSerThrLysSer-----Ser 1233
Db 288 CTCCTCTGGAGCGCTGCCCTGATTCCTTACCTCTTACCTCTTGAATAATGGTGTTCG 347

Qy 1234 AspSerAspVal-SerAspValSer-----AlaValSerArgThrSerSe 1248
Db 348 AGAAGTCAATTGAAGCTTGTGCAAAATGCCCTAGGGGTGTCTGTGTGGAGGACGCCCCC 407

Qy 1248 rAlaSerArgPheSerSerThrSerTyrMetSerValGlnSerGluArgPro----- 1265
Db 408 TGTGATGCGGAACACACAGGCTCAGATTTCATGATTCGAGCTGCCGTGAGCCCTCCACCC 467

Qy 1266 -----ArgGlyAsnArgLysIleSerValPhe-ThrSerLysMetGlnA 1280
Db 468 CGGGGACCATGTTTAAACGGGAGCCAGGTCTCTGCTCATCTGGGGCTCC----- 517

Qy 1280 snArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerLysSerGlyAspM 1300
Db 518 -----AGGAATGTGTGCGGAGCTCCAGCATTTGCGGTGAA 554

Qy 1300 etCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyT 1320
Db 555 TCTGCGGATCCAGCAGCCGCGGGC-----GGGCTGGGACCA 593

Qy 1320 hrSerGlyLysLysArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuS 1340
Db 594 CCACCGCCCAAGAAAGCGCGGAGCAGCTCGGTGCGCAAGATGCTGCCATCGTGGGCTGA 653

Qy 1340 erArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuA 1360
Db 654 CTCAGTGGAGCAAGAGCAGCACTCCAGCTTCCGCGAGCTTGAAGGGGCCCAAGAAAGCTGC 713

Qy 1360 tgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGlnMetArgAsnTpmMet 1380
Db 714 GCAGCAACATCCGCGGAGCAGGAGACAGGATCGCGTGGAGATGCGAGCCCGGTCA 773

Qy 1380 hrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyA 1400
Db 774 CAGCGCAGGCGAGCGGAGTCCACCGATGGGAGCACCACACACAGCTCCGAGCGCA 833

Qy 1400 snLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspG 1420
Db 834 CGTTTCATCTTCCCTACTACCCGCTAGGGGCTGAAAGCCAGTTCAGCGATTTCTCGATG 893

Qy 1420 lyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspI 1440
Db 894 GGCTGGGACCACTCAGATTGTGGGCGGACAGACTGGCAACACACCCATCGGAGATG 953

Qy 1440 leGlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleAlaArgG 1460
Db 954 TGCACATTGCCATCATGGACCGGAGTGGCCAGCTGGAGGTGGAGTGTGAAGCTCGGG 1013

Qy 1460 lyLeuValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuL 1480
Db 1014 GGCTGACCCCAACACAGGCTCCAAATCCTCCAGCCCACTATATCAAGTTTACCTGC 1073

Qy 1480 euAspAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeuGluP 1500
Db 1074 TGGAGATGGGCGCTCTTGGCCAGAGAACAGAACAGAGTGAAGAGAGCTGTGATC 1133

Qy 1500 roLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleI 1520
Db 1134 CCTGTGACAGCAGGCTCTGCTTTTGACAGGGGACCCCGAGGCAAGGTGCTGCAAGTGA 1193

Qy 1520 leValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleL 1540
Db 1194 TCGTCTGGGAGAGATATGGCCGATGGACCAAGTGTCTTATGGGCATGGGCCAGATCA 1253

Qy 1540 euLeuAspGlnLeuGluLeuSerAsnMetValIleGlyTrpPheLysLeuPheProProS 1560
Db 1254 TGTGACAGCTGGACCTCAGCGCGCGGTCAACCGGTGGTACAACTCTTCCCACTCT 1313

Qy 1560 erSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuG 1580
Db 1314 CCTCAGTGGAGACTCCACACTCGGATCCCTTACAGGGCGCTGTCCCACTCTTCCCTGG 1373


```
QY 414 -----AlaAsnAlaGluLeuGluAspSe 421
Db 2178 CCAGCTAGACGTGGCGGTCTACGCTCCAGAACCCAGCCAGCGTGGCGGTCTACGCTC 2237
QY 421 rArgIleSerLeuLeuArgMetAspArg-----ProSerArg----- 433
Db 2238 TAGAACCCAGCTAGACGCGGTGGTGGTCTACGCTCCAGAACACCCAGCCAGAGAGGGAG 2297
QY 434 -----GlnArgSerValSerGluArgAlaAlaMetGluAsnG1 447
Db 2298 GTCTCGGTCTAGGACACCAAGACGAGGAAGATCCCGCAGTAGAAGCTTAGTTAGAGTGG 2357
QY 447 nArgSerTyrSerMetGluArgThrArgGluAlaGlnGlyGlnSerSerTyrProGlnAr 467
Db 2358 AAGATCTCACTCT-----AGAACCTCAAAAGAGAGGACAGATCTGGCTCATCTTCAGA 2411
QY 467 gThrSerAsnHisSerProProThrProArgArgSerProIleProLeuAspArgProAs 487
Db 2412 GCGGAAACAAATCCAGAACATCTCAAGAGAGAGCAGG---TCCAATTCAGCCCA 2468
QY 487 pMetArgArgAlaAsp---SerLeuArgGlyGlnHisLeuAspProSerSerAlaVa 506
Db 2469 AATGAAGAAATCTCGCATTTCTTCAAGCGGAGCAGGTCTCTCTTCCACCGGTCCAA 2528
QY 506 lArgLysThrLys-----ArgG1 512
Db 2529 AGCAAAATCTCGTCTTGTGTTGAGCGCGACCTTTTCAGGGTCTTCCCCATGCCCTAAGCA 2588
QY 512 uLysMetGluThrMetLeuArgAsnAspSerLeuSerSerAspGln----- 527
Db 2589 AAGTCACAGACACCCAGCGGCGAGTCGTCTCGATCCTCCACCTAAAGCTAAATC 2648
QY 528 -----SerGluSerValArgProProProArg----- 537
Db 2649 TAGAACGCCACCCAGAGCGAGTCGTCCAGTCTTCTCGGCCACCTAAACAGAAATCTAA 2708
QY 538 -----ProHis-----LysSerLysLysGlyG1 545
Db 2709 GACACCATCAAGACAAAGTCATTCAGTTCATCTCTCTCATCTCTTAAAGTGAATCTGGAAC 2768
QY 545 yLysMetArgGlnValSerLeuSerSerSerGlu-----GluGluLeuAlaSerThrProG1 564
Db 2769 ACCACCGAGGCAAGGGTCCATAACAGTCCCGAGGCAATGAGCAATCTGTAAACGCCACA 2828
QY 564 uTyrThrSerCys-----AspAspValGluLeuGluSerGluSer----- 577
Db 2829 GAGACGGAGCTGTTTGAATCATCATCCTGACCTGACCTGAGTTGAATCTAGGACCCCTCTAG 2888
QY 577 ----- 577
Db 2889 ACATAGTGCTCAGGGTCTCTCTCTCTAGAGTGAATATCTAGCACACCTCCACAGACAGAG 2948
QY 578 -----ValSerGluLysGlyAs 583
Db 2949 CCATCTAGTGTATCATCTCCACACCCAAAGTGAAGGCAATATATCACCAAGACAAAG 3008
QY 583 pSerGlnLysGlyArgLysThrSerGluGlnGlyValLeuSer----- 598
Db 3009 AAGCCATCTGGCTCTCTCTTCCAACTCTCTAGTGGGTGACGTCGAGAACCACTCCACG 3068
QY 598 ----- 598
Db 3069 GCGAAGCAGATCAGTATCTCCCTGCTCCAATGTGGAATCCAGATGTTGCCAAGATACAG 3128
QY 599 -----AspSerAsnThrArgSerGluArgGlnLysLysArgMe 611
Db 3129 TCATCTGGGTCTCTCTCCACAGATACCAAGTGAACCTGAAACACCGCCCAAGACAAAG 3188
QY 611 tTyrTyrGlyGlyHisSerLeuGluGluAspLeuGluTrpSerGluPro----- 627
Db 3189 TCACCTCAGGTCTATTTCACCATACCCCAAGTAAAGGCCCAAACTCCACCGGGCCCAAG 3248
QY 628 -----GlnIleLysAspSerGlyValAspTh 636
Db 3249 TCTTCTGGATCAAGTCCACATGTCCTCCCAAGAGAGTCTAAAGACTCACTAGTTCARAG 3308
QY 636 rCys-----SerSerThrThrLeuAsnGluG1 645
Db 3309 TTGCCCTGGATCCCTCTCTCTCTGTCGAGGAGTAAATCTAGCACACCCAGCGGCGAGAG 3368
QY 645 uHis----- 646
Db 3369 CTATTTTGGTGTCTCATCTCTGCAACTGAAAGGACAATCTCAAACTCCAGACACCAAG 3428
QY 647 -----SerHisSerAspLysHis----- 652
Db 3429 ATCTCATATCTCAAGTCCAGAGTCCAGACAGAGTCAATTCAGATCCACCATCTCTGCAGAG 3488
QY 653 -----ProValThr 656
Db 3489 CAAATCTCAACATCACCTAAGGAGGTGGTCCAGGTCTTCTCATCTCCAGTCACTGAGT 3548
QY 656 rGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuL 676
Db 3549 GGCATCCAGATCTCCAATAAGACAAGATAGA--GGTGAGTCTTCACGCGAGTCTCTATGTGA 3607
QY 676 yAspGlySerValProArg----- 682
Db 3608 AATCTGGATGTCTCTCTGAGCAGACAGGTTCCAGTCTGACTCTTCTTCATATCTCTACAG 3667
QY 683 --AspSerGlyAlaMetLeuGlyLeuLysValValGly-----G 695
Db 3668 TGGACTCGAAATCTCTCTTGGGCGCAGAGTAG--ATTGGAGACTGTGTAATCAAAAGAGAAA 3726
QY 695 lYysMetThrGluSerGlyArgLeuCysAlaPheIleThrLys--ValLysLysGlySer 714
Db 3727 ATGGCTTACCCTCCAGGAGA---TGCTACTGCATCACCTCTCTAGACAGAAAGACAAA 3783
QY 715 LeuAlaAspThrValGlyHisLeuArgPro----- 724
Db 3784 TTTAGTCCCTTCCAGTACAGTAGGCTGAGTCTTCACTGGTATTCAAGACACACATT 3843
QY 725 -----GlyAspGluValLeuGluTrpAsn 732
Db 3844 AGAACCCCGCCCAAGAGAAAGTGTGTGGTGCATCTCCAGAAACAAAGAGCAAAAT 3903
QY 733 GlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleIleLeuGluSerLys 752
Db 3904 AGTGCAATTCCTACGTCAAGCCCAAGATGAAGAGTTAATGGAGGTGGTAGAGTCTGAA 3963
QY 753 ProGluProGlnValGluLeuValSerArgProIleGlyAspIleProArgIlePro 772
Db 3964 -----GAAACCGCGAGGCCAAATC-----CTGTCT 3987
QY 773 AspSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSer-----GlnLys 790
Db 3988 CATTTGTCTTCAGAACTTAAAGAAATGTCCACAAAGTAACTTTGAATCATCTCTCTGAAGTA 4047
QY 791 MetAspArgProSerLysSerValThrSerProMetSerProGlyMetLeuArg----- 808
Db 4048 GAAGAAAGCCCTGCTGTCTTGTGACTCTTGATCAGACGAGTCACAGCGCTCTTTTGGAA 4107
QY 809 -----AspValProGlnPheLeuSerGlyGlnLeuSerIleLysLeuTrp----- 823
Db 4108 GCGTAGAGTCTCTTCAATGGCCTCATCT-----TGGGTGGGCCA 4149
QY 824 ---PheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuPro 842
Db 4150 CATTTTCTCCAGAACATAAAGAACCTG-----TCTAACTCCCA 4188
QY 843 SerArgGluAsp-----GlyArgPro-----ArgAsnProTyrValLysIleTyr 857
Db 4189 CTGAGGAGAACAGCTTGGATCACCTTTAGATTTTAGAACTTCAAGCTCAGGC----- 4236
QY 858 PheLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeu 877
```

Db 4237 -----CCACTGTGTACAGAAATGAATACTGATTTTCTCTGAGTTTAAAGAAATTTG 4230
Qy 878 GluProlystPAsnGlnThrPheIleTyrSerProValHisArgGluPheArg--- 896
Db 4291 -----AATGGACCTTTCTTAA--TCAGCTGAAACAGATCCATCTTAGACAT 4337
Qy 897 GluArgMetLeuGluIleThrLeuTrpAspGln----- 907
Db 4338 GAAAGACAAATCGACAGATCTCTGGACACAGCAGTCTCTGAGTTATCCCCAGATGCAGT 4397
Qy 908 AlaArgValArgGlu-GluGluSerGluPheLeuGlyGluIleLeuLeuGluTh 927
Db 4398 GAAAAGGACGAGTGTCTTCAAAATCAG-----AGCATCTCTTC 4436
Qy 927 rAlaLeuLeuAspGluProHisTrpTyrLysLeuGlnThrHisAspValSerLe 947
Db 4437 ACCTGTGCTGATGCTGTACCCAGAACCCCTCGAGAGAAAGATAGTCTTCGATCTTC 4496
Qy 947 uPro-----LeuProArg---ProSerProTyrLeuProArgGlnLe 961
Db 4497 TCCTGAATGAAGATGTTTACCAGACTCCATCA-----AGGAGAGCAG 4544
Qy 961 uHisGlyGluSerProThrArgLeuGlnArgSerLysArgIleSerAspSerGluVa 981
Db 4545 GTCTGGGTCTCTCCAGACTTAGA----- 4569
Qy 981 lSerAspTyrAspCysGluAspGlyValValSerAspTyrArgHisAsnGlyAr 1001
Db 4570 -----GATGGTCTGGGACTCCCTCG-----AGGCACGCTCTC 4604
Qy 1001 gAspLeuGlnSerSerThrLeuSerValProGlu--GlnValMetSerSerAsnHisCy 1020
Db 4605 TGGGTCTCTCTGGAATGAAGATATACCTAGACGCCATTAGAGGAGAGCGAATG 4664
Qy 1020 sSerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPr 1040
Db 4665 TGATTTCTCCAGAACCGAAAGCTTTGCCTCAGACTCTTAGCCGAGGAGTGTCTCTCC 4724
Qy 1040 oSerAlaPro-----ProGlnArgAsnValGluGlnGlyHi 1053
Db 4725 ATCATCCCGAGAGTCAACAACAAGTGTCTTACCCCGCAGAGAA-----AG 4772
Qy 1053 sArgGlyThrArgAlaThrGlyHisTyrAsnThrIle----- 1065
Db 4773 AAGCGGTGAGATCATCAGTTGATCAGAAACTGTGCTCGACTCCCTCGGCGCAGAG 4832
Qy 1066 -SerArgMetAspArgHisArgValMetAsp-----AspHi 1077
Db 4833 AAGTCGTTCCGGATCTCTCAAGACTTGTGTGAACCCAGTGCATCCCTCAGGAAAG 4892
Qy 1077 sTyrSerSerAspArgAspArgAspCysGluAlaAlaAspArgGlnProTyr---HisAr 1096
Db 4893 AAGTGAGTCAGACTCTCTCCAGATCTTAAAGCCCAAGACAGACACCCACTTCGGCAGAG 4952
Qy 1096 gSerArgSerThrGluGlnArgProLeuLeuGluArg-----ThrThrThrArgSe 1113
Db 4953 GAGTCGGTGTGATCATCTCCGAGGTGACGCAAACTGTGCTCGACTCCCTCGCGCGAG 5012
Qy 1113 rArgSerSerGluArgProAspThr---AsnLeuMetArgSerMetProSerLeuMetTh 1132
Db 5013 TAGGTCTGTTCTCTCCCTGAGTGAAGATTAAGCCAGACAGCAGCACCAGCGCAGAG 5072
Qy 1132 rGlyArgSerAlaProProSerPro-----AlaLeuSerArgSerHisProArg----- 1148
Db 5073 TGGTTCTGATTCCTCTCTGAACTTAAAGCTCCAGCCCTCGGGCCCTTCCAGACGAGAG 5132
Qy 1149 -----ThrGlySerValG1 1153
Db 5133 CAGATCAGGTTTCATCAAGCAAGGAGAGGCCCTTCTCTCTGAAGGAGCAGCAGTACCGA 5192
Qy 1153 nThrSerProSerSerThrPro-----GlyThrGlyArgArgGly---ArgGlnLeuPr 1170
Db 5193 GTCTCTCTCTGAACATCCGCCCCCAAAATCCAGAACTCTCTCGCAGAGGTTCCAGTCTATCACC 5252

Qy 1170 oGln-----LeuProProLysGlyThrLeuGluArgSerAlaMetAs 1184
Db 5253 AGAGCCCAAGACCAAGTCTCGTACACCACTCGA-----CGTCGAGCTCTCCATC 5303
Qy 1184 pIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaGlySerAs 1204
Db 5304 ATCTCCGGAGCTAACAAAGGAAGCCGACTGTCCGTAGAAGCGCTCTGCCTCATCTCCT 5363
Qy 1204 pProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpAspPro---HisAr 1223
Db 5364 ACCA-----GAAACTCGCTCTAGAACTCCCCCAAGGCAACG 5399
Qy 1223 gGlyAlaAspThrValSerThrLysSerSerAspSerAspValSerAlaVa 1243
Db 5400 GAGAGTCTCTCAGTGTCTTCC-----CCGAGCCAGCCGCAAAATCGAGGTCTTCACGCG 5456
Qy 1243 lSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSer---ValGlnSe 1262
Db 5457 ACGCGCTCAGCTTCATCTCCAGCACTAAGACAACCTCAAGGAGAGCGCGCTCTCTCTTC 5516
Qy 1262 rGluArgProArgGly-----AsnArgLysIleSerValPheTh 1275
Db 5517 GCCAAACCTCTCGTGACTCCAGAGGTCCCGTTCCTCCGCTCAAGGAGAGAGAAAACAAGAAC 5576
Qy 1275 rSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSe 1295
Db 5577 AACCCGACGTCGAGATAGTCTGGATCTTCTCAGTCAACCTCTCGGCGAAGACAGCGGAG 5636
Qy 1295 rIleSerGlyAspMetCysSerLeuGluLysAsnAspGly-----SerGlnSe 1311
Db 5637 CCGGTCAAGGTGCGGGTTACTCGCGCGGAGGGAGGCTCTGTTATCATCTCAAGTCTC 5696
Qy 1311 rAspThrAlaValGlyAlaLeuGlyThrSerGlyLysArgArgSerIleGlyAl 1331
Db 5697 ACTGCGCGGAGGAGAAAGTTCCCGGACCTCTCTCGACGCGCAAGA----- 5742
Qy 1331 aLysMetValAlaIleValGlyLeuSerArg-----LysSerArgSerAlaSerG1 1348
Db 5743 -----GGCCCTCTCGGACACCCCCCAACCCAGTCCGAAGCGTCTTCG 5783
Qy 1348 nLeuSerGlnThrGluGlyGlyLysLysLeuArgSer-----ThrVa 1363
Db 5784 CTCAGGCACATCACCAGCCCGTGGAAACGCTCTAGATCTCGAGCCTCTCCAGCCACTCA 5843
Qy 1363 lGlnArgSer---ThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgG1 1382
Db 5844 CCGCGCATCCAGGTCCAGAACCCCTCG-----ATAAGCCGACG 5882
Qy 1382 nAlaSerArgGluSerThr 1388
Db 5883 TAGGTCCAGATCTCGAACT 5901

RESULT 5

US-09-764-891-9844/c
; Sequence 9844, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9844
; LENGTH: 24387
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-9844

Alignment Scores:

Pred. No.: 3,36e-17 Length: 24387
Score: 368.50 Matches: 354
Percent Similarity: 33.15% Conservative: 180
Best Local Similarity: 21.97% Mismatches: 548
Query Match: 4.45% Indels: 532
DB: 11 Gaps: 78

US-09-617-099B-1 (1-1590) x US-09-764-891-9844 (1-24387)

QY 193 AnSerGlySerAsnThrLeu-GlnGlnProAspGlnLysValProArgGlyLeuArgAs 212
DB 15016 AACTCGGAGCGTTCACCACTAAGTCTCC--GAGAACTTCCCGAGTCTTCTCCTC 14960
QY 212 nGluGlu-----AlaProGlnGluLysLysLysLeuHis----- 224
DB 14959 AGAGAGCAGCCACCATCCCTCACTACCAAGTTCCTCGGCATGCCAGCTTCCGCC 14900
QY 225 -GluGlnProGlnPheGlnGlyAlaProGly-----AspLeuSerValProAlaVa 241
DB 14899 AGAAAGTCTTAACCTGCTCCAGCTCCAGGGTCCACCGAGAGATTTCTTCTCCAC 14840
QY 241 lGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGlyVa 261
DB 14839 ATCTAAGATCGTCACTAGCGCGGCAAAACGGGATAATCACATTTCTCATACCCCTC 14780
QY 261 lLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSerAr 281
DB 14779 CGGTAGGATGGGAGGTCCCGTAGCCCTGCCACCGCTAAGAGAGGCGCATCTCGGTCG 14720
QY 281 sAspGlnAsnArgArgTyrgluGlnSerGluArgGluAspTyrgluSerGlnTyrgluValPr 301
DB 14719 AACCCCTACAGAGA---GGTCATCTCGATCCCGATCTCCCGAGTGGGTAGTCCAG 14663
QY 301 oSerAspGlyThrMetProArgSerProSerAspTyrgluAlaAspArgArgSerGlnArgGl 321
DB 14662 GTCTGCACAGAGTGGGGAAGATCTAGAAGCCCCCAGCGAGTGGCGCTCT---AGGTC 14606
QY 321 uProGlnPheTyrgluGluuProGlyHisLeuAsnTyrgluAsp----- 335
DB 14605 TCCTCAG-----CGACAGGCTGGTCTAGGACAGAAATACCACAGAGAGCGCAG 14555
QY 336 -----SerAsnArgArgGlyHisArgHisSerLysGluTyrgluIleValAspAspGluAs 353
DB 14554 GTCTAGTCAAGAGCGAGGAGGTCCCACTCTAGATCCCGACAGCAGTGGGTAGATC 14495
QY 353 pValGluSerArgAspGluTyrgluArgGlnArgArgGluGluGluTyrgluAlaArgTy 373
DB 14494 T---CGTCTTAGAACACACCGCCCGCGGCGAGTCCCGCTCTAGAACACCTCCAGGCG 14438
QY 373 rArgSerAppProAsnLeuAlaArgTyrgluProValLysProGln----- 387
DB 14437 GAGATCA-----CGATCCAGAACTCCACAGCGGTAGGTCTCGGTCTAGAACACC 14387
QY 388 -----ProTyrgluGluGlnMetArgIleHisAl 397
DB 14386 AGCCCGAGGGGAGGTCTCGGTCTAGAACACCTCTAGCGCGCAGATCTAGGACCCGATC 14327
QY 397 aGluVal-----SerArgAlaArgHisGluArgArgHisSerAspValSerLeu-- 413
DB 14326 ACCAGTACGACGCGAGGTCTCGTAGTAGATCACCAGCCAGGAGGTGGCAGGTCAACGCTC 14267
QY 414 -----AlaAsnAlaGluLeuGl 419
DB 14266 TAGAACCCAGCTAGAGTGGCGCTCAGCTCCAGAACCCACCGCAGAGTGGCGCGCTC 14207
QY 419 uAspSerArgIleSerLeuLeuArgMetAspArg-----ProSerArg-- 433
DB 14206 ACGTCTAGAACCCAGCTAGACGAGTGGTCTCGCTCAGCTCCAGAACACACGACCGAG 14147
QY 434 -----GlnArgSerValSerGluArgArgAlaAlaMetGl 445
DB 14146 AGGAGGTCTCGGTCTAGGACACCAAGACGAGGAGATCCGCGAGTAGAAGCTTAGT 14087

QY 445 uAsnGlnArgSerTyrgluMetGluArgThrArgGluAlaGlnGlyGlnSerSerTyrgPr 465
DB 14086 ACGTGGAGATCTCACTCT-----AGAACACCTCAAGAGAGCGAGATCTGGCTCATC 14033
QY 465 oGlnArgThrSerAsnHisSerProProThrProArgArgSerProIleProLeuAspAr 485
DB 14032 TTCAGAGCGGAAACAAATCCAGAACATCTCAAGAGAGAGCAGG---TCCAATTCAG 13976
QY 485 sProAspMetArgArgAlaAsp---SerLeuArgLysGlnHisHisLeuAspProSerSe 504
DB 13975 CCCAGAAATGAAGAAATCTCGCATTTCTTCAAGCGGAGCAGGTCTCTCTTCCACCAG 13916
QY 504 rAlaValArgLysThrLys----- 510
DB 13915 GTCCAAAGCAAAATCTCGTTGTCTTTGAGCGGACGCTTTTCAGGCTCTTCCCATGCC 13856
QY 511 -ArgLysMetGluThrMetLeuArgAsnAspSerLeuSerSerSerAspGln----- 527
DB 13855 TAAGCAAAAGTTCACAGACACACACCGCGCGAGTCGCTCTCGATCTCTCCAACTTAAAGC 13796
QY 528 -----SerGluSerValArgProProProProArg----- 537
DB 13795 TAAATCTAGACGCCACCCAGACGCGAGTCGCTCCAGTTCTTTCGCGCACCTTAAACAGAA 13736
QY 538 -----ProHis-----LysSerLysLy 543
DB 13735 ATCTAAGACACCATCAAGACAAAGTCATCTCCAGTTCATCTCTCTCATCTTAAAGTGAATC 13676
QY 543 sGlyGlyLysMetArgGlnValSerLeuSerSerSerGlu---GluGluLeuAlaSerTh 562
DB 13675 TGGAAACACACCGAGCGCAAGGTCCTAATAACAGTCCCGAGCCCAATGAGCAATCTGTAAAC 13616
QY 562 rProGluTyrgluSerCys-----AspAspValGluLeuGluSerGluSer-- 577
DB 13615 GCCACAGAGACGAGAGTGTGTTGAATCATCATCTCACCTCGCTGAGTTGAAATCTAGACCCC 13556
QY 577 ----- 577
DB 13555 TTCTAGACATAGTCTCAGGCTCCTCTCTCTCTAGAGTGAATCTAGCACACCTCCAC 13496
QY 578 -----ValSerGluLy 581
DB 13495 ACAGAGCCCATCTAGTCTCATCTCCACACCCAAAGTGAAGGCAATAATATATACCAAG 13436
QY 581 sGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSer-- 598
DB 13435 ACAAGAGACCATCTCGGCTCTCTTCTCAAGTCTCTAGTAGGCTGACGTGAGAACAC 13376
QY 598 ----- 598
DB 13375 TCCACGCGGAGCAGATCAGTATCTCCCTGCTCCAATGTGGAATCCAGATTGTGCAAG 13316
QY 599 -----AspSerAsnThrArgSerGluArgGlnLysLy 609
DB 13315 ATACAGTCAATCTCGGCTCTCTCCACAGATACCAAGTGAACCTGAAACACCGCCAG 13256
QY 609 sArgMetTyrgluGlyHisSerLeuGluGluAspLeuGluTyrgluPro----- 627
DB 13255 ACAAGTCACTCAGGCTCTATTTCACCATACCCCAAGTAAAGGCCCAACTCCACCGGG 13196
QY 628 -----GlnIleLysAspSerGlyVa 634
DB 13195 GCCAAGTCTTCTGGATCAAGTCAACATGTCCCAAGAGAGTCTAAGACTCTACTAGT 13136
QY 634 lAspThrCys-----SerSerThrThrLeuAs 643
DB 13135 TCAAGTTCCCTGGATCCCTCTCTCTGTGCGAGGAGTAAATCTTAGCACACACCGAG 13076
QY 643 nGluGluHis----- 646
DB 13075 CGAGAGCTATTTTGGTGTCTCATCTCTGCAACTGAAAGGACAACTTCAAACTTCACAGA 13016
QY 647 -----SerHisSerAspLys-His----- 652

```
Db 13015 CCACAGATCTGATATCTCAAGTCCAGAGGTGACACAGAGTCAATTCAGAAATCACCATCTCT 12956
QY 653 -----ProV 654
Db 12955 GCAGAGCAATCTCAAAACATCACTAAGGAGGTGGTCCAGTCTTCATCTCCAGTCCAC 12896
QY 654 alThrTyrGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysA 674
Db 12895 TGAGCTGGCATCCAGATCTCCATTAATACAGACAGATAGA-GGTGAGTTCTCAGCGAGTCTTA 12837
QY 674 rGleuLysAspGlySerValProArg----- 682
Db 12836 TGTGTGAATCTGGAATGCTCTCTGAGCAGACAGGAGTTCAGTCTGACTTCTTTCATATC 12777
QY 683 -----AspSerGlyAlaMetLeuGlyLysValValGly----- 694
Db 12776 CTACAGTGGACTCGAATCTCTCTGGGGCAGAGTAG-ATTGAGAGCTGCTGAATCAAAA 12718
QY 695 -----GlyLysMetThrGluSerGlyArgLeuCysAlaPheIleThrLys-ValLysLys 712
Db 12717 GAGAAATGGCTTACCCCTCAGAGGA---TGCTACTGCATCACCCTCTAGACAGAA 12661
QY 713 GlySerLeuAlaAspThrValGlyHisLeuArgPro----- 724
Db 12660 GACAAATTTAGTCCCTTTCCAGTACAGATAGGCTGAGTCTCACTGGTATTCAAGAC 12601
QY 725 -----GlyAspGluValLeuGlu 730
Db 12600 ACATTAGAACCCGCCAGAGAGAAAGTGGTGGTGCATCTCCAGAAACAAAGAG 12541
QY 731 TrpAsnGlyArgLeuGlnGlyAlaThrPheGluGluValTyrAsnIleLeuGlu 750
Db 12540 CAAATAGTCATGCTAGCTCAGCCACAGATGAGAGTTAATGGAGTGGTAGAGAG 12481
QY 751 SerLysProGluProGlnValGluLeuValSerArgProIleGlyAspIleProArg 770
Db 12480 TCTGAA-----GAACCCCGAGGCCAAATC----- 12457
QY 771 IleProAspSerThrHisAlaGlnLeuLysSerSerSerSerPheGluSer----- 798
Db 12456 CTGTCTCATTTGCTTCCAGAACTTAAAGAAATGCCCAAGTAGTAACCTTGAATCATCTCT 12397
QY 789 GlnLysMetAspArgProSerIleSerValThrSerProMetSerProGlyMetLeuArg 808
Db 12396 GAAGTAGAAGAGGCTCTGTGCTTGTACTCTTGCATCAGACCCAGTCCAGGCTTCT 12337
QY 809 -----AspValProGlnPheLeuSerGlyGlnLeuSerIleLysLeuTrp--- 823
Db 12336 TTGGAGACAGTAGAAGTCCCTTCAATGGCCTCATCT-----TGGGGT 12295
QY 824 -----PheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 12294 GGCCCAATTTTCTCCAGAACATAAAGAACTG-----TCTAAC 12256
QY 841 LeuProSerArgGluAsp-----GlyArgPro-----ArgAsnProTyrValLys 855
Db 12255 TCCCCACTCAGGAGAACACACTTTGGATCACCTTTAGAAATTTAGAACTCAGGC----- 12202
QY 856 IleTyrPheLeuProAspArgSerAspLysAsnLysArgThrLysThrValLysLys 875
Db 12201 -----CCACTTGTGTACAGAAATGAATACTGGATTCTTCTCAGGTTTAAAGAA 12154
QY 876 ThrLeuGluProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPhe 895
Db 12153 GATTTG-----AATGGACCGTTTCTTAA--TCAGCTGGAAACAGATCCATCTCT 12107
QY 896 Arg---GluArgMetLeuGluIleThrLeuTrpAspGln----- 907
Db 12106 AGACATGAAGAACAAATCGACAGATCTCTGGACACAGAGTCTTGAGTTATCCCCAGA 12047
QY 908 -----AlaArgValArgGlu-GluGluSerGluPheLeuGlyGluIleLeuIleGluLe 925
Db 11045 ATCTCTCACCA-----GAAACTCGCTCTAGAACTCCCCCAG 11045

Db 12046 TCAGTGGAAAAGCGCAGGATGTCTTCAATCAG-----AGCAT 12008
QY 925 uGluThrAlaLeuLeuAspGluProHisTrpTyrLysLeuGlnThrHisAspValSe 945
Db 12007 CTCTTCACCTGTCTGTGATGCTGTACCCAGAACACCTCGAGAGAAAGTAGTCTCTGC 11948
QY 945 rSerLeuPro-----LeuProArg-----ProSerProTyrLeuProArgAr 959
Db 11947 ATCTTCTCTGAAATGAAAGATGGTTTACCCAGAACTCCATCA-----AGGAG 11900
QY 959 gGlnLeuHisGlyLysSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSe 979
Db 11899 AAGCAGGTCTGGTCTTCTCCAGACTTAGA----- 11869
QY 979 rGluValSerAspTyrAspCysLysAspGlyValGlyValValSerAspTyrArgHisAs 999
Db 11868 -----GATGGTCTGGGACTCCCTCG-----AGGCACAG 11840
QY 999 nGlyArgAspLeuGlnSerSerThrLeuSerValProGlu---GlnValMetSerSerAs 1018
Db 11839 CTTGTCTGGTCTCTCTCTGGAATGAAGATATACCTAGAACGCCATCTAGAGGGAGAG 11780
QY 1018 nHisCysSerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTr 1038
Db 11779 CGAATGTGATTCTTCCCGAAGCCGAAAGCTTTTGGCTCAGACTCTTAGGCCGAGGAGTGC 11720
QY 1038 pSerProSerAlaPro-----ProProGlnArgAsnValGluGlu 1051
Db 11719 TTCTCATCATCTCCAGAGCTCAACACAAAGTGTCTTACCCCCAGAGAGAA----- 11668
QY 1051 nGlyHisArgGlyThrArgAlaThrGlyHisTyrAsnThrIle----- 1065
Db 11667 ---AGAAGCGGTCCAGATCATCATCTGATCAGAAAAGTGGTCCGACTCCCTCGG 11612
QY 1066 ---SerArgMetAspArgHisArgValMetAsp----- 1075
Db 11611 GCAGAGAGTCTGTCGGGATCTCTCAAGAACTTGATGAAAACCCAGTGCATCCCTCA 11552
QY 1076 -AspHisTyrSerSerAspArgAspCysGluAlaAlaAspArgGlnProTyr--- 1094
Db 11551 GCAAGAAAGTGAAGTCAAGCTCTTCTCCAGATTCTAAAGCCAGACACGAAACCCATTCG 11492
QY 1095 -HisArgSerArgSerThrGluGlnArgProLeuLeuGluArg-----ThrThrTh 1111
Db 11491 GCAGAGAGTCTGTCGGATCATCTCCAGAGGTGGACAGCAATCTCGATATCCCTCG 11432
QY 1111 rArgSerArgSerSerGluArgProAspThr---AsnLeuMetArgSerMetProSerLe 1130
Db 11431 GCGCAGTAGTCTGTGTTCTTCCCTCAGTGAAGATAAGCAAGAGCAGCACCCAGGGC 11372
QY 1130 uMetThrGlyArgSerAlaProProSerPro-----AlaLeuSerArgSerHisProAr 1148
Db 11371 ACAGATGTTCTGATTCTCTCTGAACTAAAGCTCCAGCCCTCGGGCCCTCCCGAG 11312
QY 1148 g-----ThrGlySe 1151
Db 11311 ACGAACAGATCAGGTTTCATCAAGCAAGCAGAGGCCCTTCTCTCGAAGGAGCAGCAG 11252
QY 1151 rValGlnThrSerProSerSerThrPro-----GlyThrGlyArgArgGly---ArgG 1168
Db 11251 TACCGAGTCTCTCTGAAATCCGCCAAATCCAGAACTGCTCCGACAGAGTTCAGGTC 11192
QY 1168 nLeuProGln-----LeuProProLysGlyThrLeuGluArgSerAl 1182
Db 11191 ATCCAGAGAGCCCAAGACCAAGTCTCGTACACCCCTCGA-----CGTCGACGCTC 11141
QY 1182 aMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaG 1202
Db 11140 TCGATCATCTCCGAGCTTAACAGAGAGGCCAGACTGTCCGTTAGAGCGCTCTGCCTC 11081
QY 1202 ySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpAspPro-- 1221
Db 11080 ATCTCTCACCA-----GAAACTCGCTCTAGAACTCCCCCAG 11045
```


Qy 271 erAspArgLysArgSer-----ProSerValSerArgAspGlnA 284
Db 8205150 CCGCGCTCGCGCTGACAGCTGATGCGCCCTCACCCTCAGCGCGGAGCGCGC 8205091
Qy 284 snArgArgTyrGluGlnSerGluGlnArgGluAspTyrSerGlnTyrValProSerAspG 304
Db 8205090 GTCCAGCGCGCGCGCGCGCTCGG-----CCGCGACGAC 8205055
Qy 304 lyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArgGluProGlnP 324
Db 8205054 GACGAGCGCGCGCGCGCGCGGAGCGGTGCGCTGACGCGCGCGCGGAGTCTCTCGGC 8204995
Qy 324 heTyrGluGluProGlyHisLeuAsnTyrArg-----AspSerAsnArgArgGlyH 341
Db 8204994 CCACTGGCGCGCGCGCGCGCTCCACACGCGCGCGCGGAGAACCCCGCGCGCGGAGGGA 8204935
Qy 341 isArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyrG 361
Db 8204934 CTTCCACAACTCCG-----C 8204920
Qy 361 luArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAlaA 381
Db 8204919 CCGCGCGGTACCGCGCGCGCGAGCGCGGTTCGGCGCGCGGAGCGCGCGCGCGCTC 8204860
Qy 381 rgTyrProValLysProGlnPro-----TyrGluGluGlnMetArgIleHisA 397
Db 8204859 GCGCGTCTTCCGCGCTCGCGCGCGCGCGCTGCGCGCGGAGCGCGCGCGCGCGCACATC 8204800
Qy 397 laGluValSerArgAlaArgHisGluArgHisSerAspValSerLeuAlaAsnAlaG 417
Db 8204799 GCGGAGGCTCTCCGCGGTACCGCGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 8204740
Qy 417 luLeuGluAspSerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerV 437
Db 8204739 GAGCACGG---GCTCGGGGACACGACTGCGCGAGCGCGCGCGCGCGCGCGCGCGCTC 8204683
Qy 437 alSerGluArgArgAlaMetGluAsnGlnArgSerTyrSer----- 451
Db 8204682 GGACTCGCGGACCGCGCTCGAGGAGCGCGTCTGATCAGTCAGCGCGCGCGCTCCGCGTGC 8204623
Qy 452 -----MetGluArgThrArgGluAlaGlnGlnGlnSer----- 462
Db 8204622 GTCCCGCACGCTGCGCGCGCGCGCTCGGACCGCACGAGCGCGCTCCTCCGCGTGCAGC 8204563
Qy 463 --SerTyrProGlnArgThrSerAsnHisSer----- 472
Db 8204562 GACGCGACCGCATTCCTCGGCGCACGCGATAGTCGAGTACCGATCAAGATGAAGTACGCC 8204503
Qy 473 -----ProProThrProArgArg----- 478
Db 8204502 CTTGTCCAGCACGACGCGCGCGCGCACTCTTGGCCAGCGCGCTCCCGCCACGT 8204443
Qy 479 -----SerProIleProLeuAspArgPro---AspMetArgArgAlaA 492
Db 8204442 CCGCGGCTTGCGCTTGGCGCACTTGGCGCTTCCGCGCGCGCTCAGCGATCAACGAGCTG 8204383
Qy 492 spSerLeuArgLysGlnHisLeuAspProSerSerAlaValArgLysThrLysArgG 512
Db 8204382 CTCGCGCGCGCGCTTCTTGGCACCGCGCGCGCGCGCTGCGCTGCTGCTCACACGAGC 8204323
Qy 512 luLysMetGluThrMetLeu-----ArgAsnAspSerLeuSerSerAspG 527
Db 8204322 ATTCTTACCAACCCCACTGACACGCGCGCGCGCGCGCGCGCGCTCGGTGCGGATCAC 8204263
Qy 527 lnSerGluSerValArgProProPro-----ProArgProHisLysSerLysLysG 545
Db 8204262 GCACGCGCGAGGCGGCTT---CCGCGCATGACCGCTCGCGCATACCGGTCCAGGACCGCGCA 8204204
Qy 545 lyLysMetArgGlnValSerLeuSerSerGluGluGluLeuAlaSerThrProGlu- 564
Db 8204203 GCAGGATTCGAACCTCGGAG-----GGAGACACCCCGAGAG 8204168
Qy 565 TyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGluLysGlyAspSer 584

Db 8204167 TACCGCACCGCTCGCACCGCGCGCGCTCGTCCGCGCTCCCGAGATCGATCCGCGC----- 8204114
Qy 585 GlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSerAsnThrArgSer 604
Db 8204113 ---CACGGCTCGTCCCAAGTTTCCAGCAC-----GCCAAGCACCGCTCC 8204072
Qy 605 GluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGluAspLeuGluTyr 624
Db 8204071 ACCGAATATCCAGAGCTCTCCGACCGCGGAGATCACAGCGTACGAGACGCCACTTTGG 8204012
Qy 625 SerGluProGlnLysAsp-----SerGly-----ValAspThrCys 637
Db 8204011 TCCGCCAGGCATTTGCTGCTACCGATCCGCGCGCGCGCACCGCGCTCGACACGCGC 8203952
Qy 638 SerSerThrThrLeuAsnGlu-----Glu 645
Db 8203951 GAGCCCGCGCACCCCGAGGGAATCCCGGCTCCGAAACAGAACGTCAGCGCGTTACAGCCC 8203892
Qy 646 HisSerHisSerAspLys---HisProValThrTyrGlnProSerLysAspGlyAspArg 664
Db 8203891 CACAGCGCTCCGCGCGCGCTCCACCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 8203835
Qy 665 LeuIleGlyArgIleLeuLeuAsnLysArgLeu-----LysAspGlySer 679
Db 8203834 CTCGCGCGCGAGTGGCGCTACCGCGCTCTGAGTAGATCGGGCGGAGCAGGTGCGAG 8203775
Qy 680 ValProArgAspSerGlyAla-----Met 687
Db 8203774 GTCGCGGATGATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTCGATCTT 8203715
Qy 688 LeuGlyLeuLysValValGly-----GlyLysMetThrGluSerGlyArgLeuCysAla 705
Db 8203714 GTCGCGCTGACCTTGGCGTGGCGAGGTCTCGACGACCGCGCGCGCGCGCTC----- 8203661
Qy 706 PheIleThrLysValLysLysGlySerLeuAlaAspThrValGlyHisLeu----- 722
Db 8203660 -----GGCTTCCGATCGGTACGCCACCTGGACCTGCGCA 8203625
Qy 723 -----ArgProGlyAspGluValLeuGluTyrPasnGlyArgLeuLeuGlnGly 738
Db 8203624 GCGGAGCGCGCACCGC 8203565
Qy 739 AlaThrPheGluGluValTyrAsnIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 753
Db 8203564 GGAGCGGCTC---GACCTTGGACCGGCTTTCGCGGAGAGCGACCGCGCGCGCGCGCGC 8203508
Qy 754 ---GluProGlnValGluLeuValVal----- 761
Db 8203507 CATGCGCGCTGATGTCGATGATGATCTTTCGCGCGCGGTGAGACCGCGCTCACCATCGG 8203448
Qy 762 -----SerArgProIleGlyAspIleProArgIleProAspSerThrHisAl 777
Db 8203447 GCGCGCGATCTCGAAGCGCGCGCGGTGGGTGATGACGAGAGCGGTAGCGCTCGTCTCCAG 8203388
Qy 777 aGlnLeuGluSerSerSerSerSerPhe----- 786
Db 8203387 CTTGATGCGCTGTCGAGGAGCGCTTCACTCCGCTCCACCGACCGCGCGCGCGCGCGC 8203328
Qy 787 -----GluSerGlnLysMetAspArgProSerIleSerValThrSerProMe 802
Db 8203327 GCGGCGGAGGAGGAGTCCAGGTGTCGAGGCGGTGTCGAGAGGAGGAGGAGGAGGAGGAG 8203268
Qy 802 tSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIleLysLe 822
Db 8203267 GTCCAGACGACCGCGC-----TTGTCGCGCTCGTCTGATGATGATGATGATGATGATGAT 8203226
Qy 822 uThrPheAspLysValGlyHisGlnLeu-----IleValThrIleLeuLeuGlyAlaLys----- 839
Db 8203225 CTGGGTCTTTCGTCGCGGAGCGAGGTACGGGATCGTCCGCTTCTTGGGACCTCGGACAG 8203166
Qy 840 -----AspLeup 842

Db 8203165 GCGGTGCGACACAGCGTGGCGAGGTGGATCGGACGGCATCAGCGTGGCGTCTCGTC 8203106
Qy 842 roSerArgGlu-----AspGlyArgProArgAsnProTyrValIleTyrP 858
Db 8203105 CGTCGGTAGCGGAACATCAGGCGCTCGTGGCGCGCGCTCGCGT----- 8203059
Qy 858 heLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValIleLysThrLeuG 878
Db 8203058 -----CCAGCTCGTCTCGTCCGCTCGACCGCTCTCGTACGCGTGTGCACACCGCTG 8203004
Qy 878 luProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArg-Glu 897
Db 8203003 CGCGATGTCGGGGACTCGGACCGGATCGACACCGACCGCGCAGGAAGCTCGTCTGAA 8202944
Qy 898 ArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGluPhe 917
Db 8202943 GCGCTTCTTCGAGAGTCGTAGCGCAT-----CTCCAGGATCTTGTGCGGACCGAG 8202893
Qy 918 LeuGlyGluIleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTyr 937
Db 8202892 CTGGCGGATCGGTGCGTACGCTTGTGTGTGACCTCGCGGC----- 8202851
Qy 938 LysLeuGlnThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuPro 957
Db 8202850 -----CAGTGCACAGCGCGTCTG----- 8202830
Qy 958 ArgArgGlnLeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSer 977
Db 8202829 GATCAACGCTCCACCGCGAC-----CCGGATGTCGGGTCTCTACGCGAAGCGCATCAAG 8202773
Qy 978 AspSerGlu-----ValSerAspTyrAspCysGluAspGlyValGlyVal-Valse 994
Db 8202772 AATGTGTCTCGTATCTGTGTACGCGATCTT-----GTCCGGGTGCGCTCGCTCAC 8202722
Qy 994 rAspTyrArgHisAsnGlyArg----- 1001
Db 8202721 GGACTCCGAGGTGAACAGCGACCGGACACACAGCTCCCTGGGGTTCAGCGGTCTCTGG 8202662
Qy 1002 ----AspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCy 1020
Db 8202661 CTGATCATTTGGCGAGCGGAGCGAGGTGCGCCCGCGCTCGTC----- 8202619
Qy 1020 sSerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPr 1040
Db 8202618 -----CGAGAACAGTTTATCGGTTCGCACTCG-----CC 8202590
Qy 1040 oSerAlaPro----- 1044
Db 8202589 ATCCGCCCCCTGTCTCGCTCTCGGAGGGCTGTGACCTGCGGCGACGGGATCTGCGCC 8202530
Qy 1044 oProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGlyHisTyrAsnTh 1064
Db 8202529 AATGCCACGGGTCTCCGCGAGCTCACACACCGCTTTCGAGCGATTTCGAGACGGCC 8202470
Qy 1064 rIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSerAspArgAspAr 1084
Db 8202469 TTTGGCGAATGGGACCGTCAC-----CCGAGCGGGGCGACGACCATCCCA 8202422
Qy 1084 gAspCysGluAla-----AlaAspArgGlnProTyrHisArgSerArgSerThrGluGlnAr 1103
Db 8202421 CACGGTCTCGGCCAAGCGCTCTTGGCGCGGTACCGGACCGGATCTCGTTCGCTCGCG 8202362
Qy 1103 gProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProAspThrAsnLe 1123
Db 8202361 GCG-----AGCACCGCGCTCTGCTCTCTCGAG-----CCGAAAGT 8202323
Qy 1123 uMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerProAlaLeuSe 1143
Db 8202322 CTTGCGCTCCCGCACCTCGTTCCACGAGGAGTTCGACCGCTTGGCTGCCAGCTTCGT 8202263
Qy 1143 rArgSerHisProArgThrGlySerValGlnThrSer---ProSerSerThr---ProGl 1161
Db 8202262 ACGGCGGTGGCGAGGACGTCTGCTCGCGGGGAGCGGAGCCGACGACCATCTGACCGGG 8202203

Qy 1161 yThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArgSe 1181
Db 8202202 GCGCGCGCGGTCTC-GGCCGAGATCTC-----CGGA 8202174
Qy 1181 rAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLys-----TyrLysG 1199
Db 8202173 GAATGTCGGATTACGACACAGACGATCGTTCGGCTCTCTGGCGTCTCTCTCTTGA 8202114
Qy 1199 nValaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTr 1219
Db 8202113 TCTTTCGGTTCGCGTACGTGCG-----CGACGGAAGTCCGCCACGCGCG 8202069
Qy 1219 pAspProHisArgGlyAla----- 1225
Db 8202068 CGCCATCACCGCGCTCGGCTCCGAGCGCGCTTCAGCACCGCTTCGCGAGTTGCA 8202009
Qy 1226 -----AspThrValSerThrLysSer-SerAspSerAspValSerAspV 1240
Db 8202008 CGCGCGTCCGACCTGGACAGCTGCACACCGCAGTTCGGCGAGCGGTTCGCGG 8201949
Qy 1240 alSerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerV 1260
Db 8201948 CGACGAGGTGACCGCGCGCGCGCGCGTCTCGCGAGGCGGTACCCCTGCT 8201889
Qy 1260 alGlnSerGluArg---ProArg-----GlyAsnArgLysIleSerV 1273
Db 8201888 TGCCGAGGAGCGGTGTCGAGAGAGCGGACCGGTCGAGCGGTTCGCGCGTACCGCGG 8201829
Qy 1273 alPheThr-----SerL 1277
Db 8201828 CGCTGACGACACAGTTCGCGCGCGGAGTTCGCGCGCGCGCGCGCGGACGACCC 8201769
Qy 1277 yMetGlnAsnArgGlnMetGlyVal---SerGlyLysAsn---LeuThrLysSerThrS 1295
Db 8201768 GCGCGCAGACCTCGAAGATCTCCGTGGGTTCGGGCGAGCGCGCTTCGCGGTTCGACGC 8201709
Qy 1295 erIleSer-----GlyAspMetCysSerLeuGluLysAsnAspGly----- 1308
Db 8201708 CGGTGACGCGCGCCACCGCGGCTCGATGACGAGCGCGCGCGCGCGCGCGGACG 8201649
Qy 1309 -----SerGlnSerAspThrAlaValGlyAlaLeuGlyThrS 1321
Db 8201648 CATTCTCTCGGTGCGCGGTGTTCCACATCTCGTGTGATGCGGTGCGAGACGA 8201589
Qy 1321 erGlyLysLysArgArgSerSerIleGly----- 1330
Db 8201588 CCGGACAGCGTCCGTGACGAGCGTTCGTGACGAGTTCGTGCGGAGCGCGTGGCGCG 8201529
Qy 1331 -----AlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnL 1349
Db 8201528 CCTTGGCGAGCATGTCGGCGGTGGCGGG---GCGACGACGACGAGTTCGGGTGCTGCGC 8201472
Qy 1349 euSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerThrGluT 1369
Db 8201471 CGATCGGACGTGCGGACCTCG-----TGACGCTGTCGCCAGACCTCGGTGAGA 8201421
Qy 1369 hrGlyLeu----- 1371
Db 8201420 CCGGTTCGCGAGAGCGGACAGTTCGGCGCGCGCGGAGTTCGACGCGCGGCGCG 8201361
Qy 1372 -----AlaValGluMetArgAsnTrp----- 1378
Db 8201360 TGGGACGACGCGGAGTGTCTCCGATCTCGTCCGCTCCGCGACGAGCTCACAGGCT 8201301
Qy 1379 -----MetThrArgGlnAlaSerArgLysSerThrArgLysSerThrAspGlySerM 1392
Db 8201300 TGAGCGCGCATGCGCCACTGATCCCGACGACGACCTTCGGTTCGCGGTTCACCGGCTC 8201241
Qy 1392 etAsnSerTyrSerSerGluGlyAsnLeuIle----- 1402
Db 8201240 CCCACACTCGTGAATCTACGGCTCGATGCTCGAATCGCGAGGGGAGCCACCGCGCG 8201181

```
QY 1403 -----PheProGlyValArgLeuAlaSerAspSerG 1413
Db 8201180 CCGCGCACACACAGCGCCGACGTCGCGCGCTGAGATGAGCGAACAATGAGC 8201121
QY 1413 InPheSerAspPhe-----LeuAspGlyLeuGlyProAlaGlnLeuValGlyA 1429
Db 8201120 GAACGAACGC-CTAGTCGCGCGCGCGCTCGACGCGCTCGGA-----GGTC 8201077
QY 1429 rGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLysG 1449
Db 8201076 AGCAGACCGCGCTGATCTCGCGACGGCGAT----- 8201045
QY 1449 lYglnLeuGluValGluIleAlaArgAlaArgGlyLeuValLysProGlySerLysT 1469
Db 8201044 -----CGAGAGCGG----- 8201036
QY 1469 hrLeuProAlaProTyrValLysValTyrLeuLeuAspGlnGlyValCysIleAlaLysL 1489
Db 8201035 -----CTTCTCGTGGACGTGGGTGTC-----GACG 8201011
QY 1489 ysLysThrLysValAlaArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSerPheG 1509
Db 8201010 AGGGGACCGACGACTCGAGGAGCGCTCACCGAGCTG----- 8200973
QY 1509 luGluSerProGlnGlyArgValLeuGlnIleIleValTpgLysAspTyrGlyArgMetA 1529
Db 8200972 -----CGAGTAGTAGCGGTGTGATCTGACGGCGCGCTTGGCGCGGTAG 8200930
QY 1529 spHisLysSerPheMetGlyValAlaGlnIleLeuLeuLeuAspGluLeu 1544
Db 8200929 ATCAGAGGCTGTACTTCGAGTCACTGCG-----CTCGAGGAGCTC 8200889
```

RESULT 7

```
US-09-814-353-20083
; Sequence 20083, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20083
; LENGTH: 9169
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-20083
```

```
Alignment Scores:
Pred. No.: 1,45e-12 Length: 9169
Score: 300.50 Matches: 369
Percent Similarity: 31.20% Conservative: 217
Best Local Similarity: 19.65% Mismatches: 687
Query Match: 3.63% Indels: 606
DB: 12 Gaps: 78
```

```
US-09-617-099b-1 (1-1590) x US-09-814-353-20083 (1-9169)
QY 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
Db 2193 ATGTCAACCCCAAGCTGCTCAAGGAGAGACG-GTGAAAAAGGAAACAAAGGTAAAGCCT 2251
QY 21 ProGlnProGluMetProAsp-----LeuSerHisLeuThrGluGluGluArgLys 38
Db 2252 GAAGACAAAGAGAGAGAGAAAGAAAGAAAGAAAGAAAGAGTGTCTAAAGAGGAGGACAA 2311
QY 39 lleileLeuAlaValMetAspArgGlnLysLysGluGluGluLysGlnSerValLeu 58
Db 2312 ACACCTATCAAGAGAGGAGAAACCCAAAAAGGAGAGGTGAAAAAGAGAGTCAAAAAA 2371
QY 59 lylsIleLysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThr 78
Db 2372 GAGATCAAGAAAGAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 2407
QY 79 GluLeuValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGlnLysGlu-ProGln 98
Db 2408 GAGGTTAAGAAAGAAACACCCGCAAGGAGAGTCAAGAGAAAGTTAAGAGAGGAGAGAA 2467
QY 98 nThrLysLeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGln 118
Db 2468 GAAGGAGTGAAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTCCCTAAAG 2527
QY 118 userGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 138
Db 2528 CCAAGAGAAATCATCTACTCTCTCTGTAAGCAAAAAAACCCAGC---TGCTTTAAAGC 2584
QY 138 rLysPheAlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCys----- 156
Db 2585 CAAGAGTACC-----CAAGAGAGAGAGTCTGTCAAGAAAGATTTCTGTGCTG 2632
QY 157 -----AlaArgCys 159
Db 2633 CCGAAAGCCAAAGGAGAGGAGGAAATAAAGTCAATTAAGAGAGGAGGAGGAGGAG 2692
QY 159 sGlyArgValSerLeuArgSer----- 167
Db 2693 AGGCTGTCTGCTGAGTGTGCGCAGTGGAGCCACACAGCAGCTGTCTATGGCGGAGCTG 2752
QY 168 -AsnLysValMetTyrValCysAsnLeuCysArgLys-GlnGln----- 181
Db 2753 GAATAGCAGCCATGCGCTGCCCAAGAACTCGAAGCTGAGAGTGCCTTATGTCTATCTC 2812
QY 182 --GluIleLeuThrLysSerGlyAlaTyrPheTyrAsnSerGlySerAsnThrLeuGlnG 201
Db 2813 CTGAGGATCTAACCAAGGAC-----TTTGAAGAGTTAAAGG 2848
QY 201 InProAspGlnLysValProArgGlyLeuArgAsnGlnGluAlaProGlnGluLysA 221
Db 2849 CTGAAGAGGTGCGATGTAAACAAAGGACATCAAGCTCAGCTGAGCTAATCGAAGACGAAG 2908
QY 221 lalysLeuHisGlnGlnProGlnPheGlnGlyAlaProGlyAspLeuSerVal----- 238
Db 2909 AGAAACTGAAGGAAACT-----GAGCCAGTCCGAGGCTACGCTCATCCAGA 2953
QY 239 -----ProAlaValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysA 257
Db 2954 AGGAGAGAGAGTCCAAAGGTCTGCGAGTCCCTGATGAGGAGAAATCATACACATG 3013
QY 257 snGlySerGlyValLysHisGlnIleAlaSerAspMet---ProSerAspArgLysArgS 276
Db 3014 AAGGGGAGGCGGAATGTGAACACAGACACCTGAGGAGCTGGAGCCCTCGAAGAGCGAGGAG 3073
QY 276 erProSerValSerArg-----AspGlnAsnArgArgTyrGlnGlnSerGluArgG 294
Db 3074 TAGACGACATTTAAAAATTTGAAGATGAAGAGCGCGTGTTCAGAGAAATCTTCAGAGACTG 3133
QY 294 luAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAspTyrA 314
Db -----
```

Db 3134 GAGACTATGAGAG----- 3147
Qy 314 laAspArgSerGlnArgGluProGlnPheTyrGluGluProGlyHisLeuAsnTyrA 334
Db 3148 -----AAGCAGAACTGAGGAGGCTGAGGACCA----- 3177
Qy 334 rgAspSerAsnArgArgGlyHisArgHis-----SerLysGluTyrIleV 349
Db 3178 -----GAAGAGGATGGGAGGAAACACGATATGTGTGAGCGCTCCAAAGCAGACGCCCA 3229
Qy 349 alaAspGluAsp-----ValGluSerArgAspGluTyrGluArgGlnArgArgGluG 367
Db 3230 CTGAGGATGAGAAAGTCCCAAGCGGAGGCTGATGATATCATCAGGGAGAGAGGAGT 3289
Qy 367 luGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAlaArgTyrProValLysProG 387
Db 3290 CTGTGGCCAGTGGGATGACCGAGCGGAGGACATGATGATGAGGCCATTGAGAAGAGGAG 3349
Qy 387 lnProTyrGluGln-MetArgIleHisAlaGluValSerArgAlaArgHisGluArg 406
Db 3350 AG-----GCTGAACAATCTGAAGCAGGAGGCTGATGAGGAGGACAAAGCTGAAGATGCC 3403
Qy 407 ArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeu--- 425
Db 3404 AGAGAGGAGGAATATGACCCCGGAAAAAATGAAGCTGAAGACTATGTGATGGCTGTGTC 3463
Qy 425 ----- 425
Db 3464 GACAAGGCTGCAGAGGCTGGTGTGCCGAGGAGCAGTATGATTCTCCACACACCAACC 3523
Qy 426 -----LeuArgMetAspArgProSerArgGlnArgSerValSerGluArgArgAlaA 443
Db 3524 AAGCAACTAGGAGGCCAGTCTCTGGCCGAGAACCTGCATCTTCAATTCTATGATGAGACT 3583
Qy 444 Met-----GluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGln 459
Db 3584 TTACTTGAGGCTCAGAGGAGGCCACCGCTCTGTGATGAGGAGATCGAGAGACCGAG 3643
Qy 460 GlyGln-----SerSerTyrProGlnArgThrSerAsnHisSerProPro 474
Db 3644 CTGAGGAATTCACTGCCACCTCTGGCTACACTCAGTCTACTATTGAGATATCC----- 3697
Qy 475 ThrProArgArgSerProIleProLeuAspArgProAspMetArgArgAlaAspSerLeu 494
Db 3698 -----AGTGAGCCCAACCCCATGGATGAG----- 3721
Qy 495 ArgLysGlnHisLeuAspProSerSerAlaValArgLysThrLysArgGluLysMet 514
Db 3722 -----ATGCTCTACCCCTCGAGACGTGATG 3745
Qy 515 GluThrMetLeuArgAsnAspSerLeuSerSerAspGlnSerGluSerValArgProPro 534
Db 3746 AGTGATGACCAACATGAGAGAGCGGAGTCCCTCTCAGGAATTC----- 3793
Qy 535 ProProArgProHisLeuSerLysGlyGlyLysMetArgGlnValSerLeuSerSer 554
Db 3794 -----GTAAATATCAACCAAA 3808
Qy 555 SerGluGluGluLeuAlaSerThrProGluTyrThrSerCysAspAspVal---GluLeu 573
Db 3809 TATGAATCTTCATTGATTCT---CAGGAATACTCTAAACCTGCTGATGTTACACGGCTC 3865
Qy 574 GluSerGluSerValSerGluLysGlyAspSerGlnLysGlyLysArgLysThrSerGlu 593
Db 3866 AACGGATTCTCAGGATCAAAACAGATGCCACTGATGGCAAGGATTACAATGCTTCA 3925
Qy 594 GlnGlyValLeuSerAspSerAsnThrArgSerGluArgGlnLysLysArgMetTyrTyr 613
Db 3926 GCCTCTTACCATATCACCCCTCT----- 3949
Qy 614 GlyGlyHisSerLeuGluGluAspLeuGluTyrSerGluProGlnIleLysAspSerGly 633
Db 3950 -----TCCATGGAGGAGAC---AAATTACAGCAGATCTGCTTTACGTGATGCT--- 3994

Qy 634 ValAspThrCysSer-----SerThrThrLeuAsnGluGluHisSerHisSer 649
Db 3995 -----TACTGCTCTGAAGTGAAGCAGCACCTTTGGACATCAAGATAGCATCTCA 4048
Qy 650 AspLysHisProValThrTrpGlnProSerLysAspGlyAspArgLeuIleGlyArgIle 669
Db 4049 GCTGTTTCAAGTGAAAGGTCAGCCCATCGAAG----- 4081
Qy 670 LeuLeuAsnLysArgLeuLysAspGlySerValProArgAspSerGlyAlaMetLeuGly 689
Db 4082 -----AGCCGCTCCTGAGTCCATCTCCACCATCACCTTAGAAAAAGACCCCTGGGT 4135
Qy 690 LeuLysValValGlyGlyLysMetThrGluSerGlyArgLeuCysAlaPheIleThrLys 709
Db 4136 GAACGTAGTGTGAACCTTCTCTGACGCCCAAT----- 4168
Qy 710 ValLysLysGlySerLeuAlaAspThrValGlyHisLeuArgProGlyAspGluValLeu 729
Db 4168 ----- 4168
Qy 730 GluTrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluValTyrAsnIleLeu 749
Db 4168 ----- 4168
Qy 750 GluSerLysProGluProGlnValGluLeuValValSerArgProIleGlyAspIlePro 769
Db 4169 -----GAGATTAAAGTCTCTCGAGGCGAGAGTAGCCCGC 4204
Qy 770 ArgIleProAspSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGln 789
Db 4205 GTGCTCTCTGAGGTGACCCAGAGTAGTGTGAGAACATTGTCTATCTCTGAGACAG 4264
Qy 790 LysMetAsp-----ArgProSerIleSerValThrSerProMetSerProGlyMetLeu 807
Db 4265 ACTCTGGAAGTGGTGTCCCATCTCAGTCCGTGACT-----GGCAGTGCT 4309
Qy 808 ArgAspValProGlnPheLeuSerGlyGlnLeuSerIleLysLeuTrpPheAspLysVal 827
Db 4310 GGTCAACACCTTACTATCAATCTCTACT-----GACGAGAAATCC 4351
Qy 828 GlyHisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuProSerArgGluAspGly 847
Db 4352 AGTCAT-----CTCCCTACAGAGTCATTGAA 4378
Qy 848 ArgProArgAsnProTyrValLysIleTyrPhe---LeuProAspArgSerAspLysAsn 866
Db 4379 AAACCA-----CCAGCAGTTCAGTGAGTTTGAATTCAGTGATGCCAAAGATGAGAA 4432
Qy 867 LysArgArgThr-----LysThrValLys 874
Db 4433 GAAAGGGCTTCAGTAAGCCCATGATGAGCCCGTCTGACTCAGAGTCTCTTATTGAA 4492
Qy 875 LysThrLeuGluProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGlu 894
Db 4493 AAAGTTTGTCTCTTTACGCGCCCGCCCTCATTTGATCCGAGTCTGCTATGAAAGT 4552
Qy 895 PheArgGluArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValArgGluGlu 914
Db 4553 TTT-----CTAAGTCTGATGACAGGCTTCTGGCAGAGGTGCCGAA 4594
Qy 915 SerGluPhe-----LeuGlyGluIleLeu 922
Db 4595 AGTCTCTTTGAAGAAAGAGAGTGGAAACAAGGCTCTCCAGACCAAGTAAGTCCAGTTCT 4654
Qy 923 IleGluLeuGluThrAlaLeuLeuAspAsp----- 932
Db 4655 GAATGACTTCTACTAGTCTTTTACAGACAAACAGGAGGAGAAAGACAGACTTTGCA 4714
Qy 933 -----GluProHisTrpTyrLysLeuGlnThrHisAspValSerSerLeuProLeu 949
Db 4715 CCAATAAAGAGACTTTGGCCCAAGAAAGAAACTGATGATGTTGAAGCCATGAGT--- 4771

```
Qy 950 ProArgProSerProTyrLeuProArgArgGlnLeuHisGlyGluSerProThrArg--- 968
Db : : : : :
4772 TCTCAACAGCACTGGCTCTCGATGAAGAAATAGGATGTTCTCCACACAATA 4831
Qy 969 -----ArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
Db : : : : :
4832 GATGTCAGTCAGTTGGATCTTTAAAGAGACACTAAGATGTCCATTTCTGAAGGTACT 4891
Qy 981 ValSerAspTyrAspCysGluasp---GlyValGlyValValSerAsp---TyrArgHis 998
Db : : : : :
4892 GTCTCAGCAAGTCACTGCTCTGTTGATGAGGCGTAGCAGAGACACGATCTCTCAT 4951
Qy 999 AsnGlyArgAspLeuGlnSerSerThrLeuSerVal-----ProGluGln 1013
Db : : : : :
4952 ATGAGGGTGTGGCTCAGTGTCCACAGCCTCAGTGGCTACGAGTCAATTTCCAGAGCCA 5011
Qy 1014 ValMetSerSerAsnHisCysSerProSer-----GlySerProHisArg 1028
Db : : : : :
5012 -----ACACAGATGATGTCTCCATCTCTGCATCTGCTGAGTTTGGCTCCCAATTC 5065
Qy 1029 ValAspVal-----IleGlyArgThr 1035
Db : : : : :
5066 ACAGAGTAGATGACTCCCTTTCAGTGTCTGTGTGGAAACACCTACCACATTCAGGAA 5125
Qy 1036 ArgSerTrpSerProSer-----AlaProPro--- 1044
Db : : : : :
5126 ACAGAAATGTCTCCATCTAAAGAGAAATGCCAAGACCGATGTCAATTTCTCCACGAT 5185
Qy 1045 -----ProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAla 1058
Db : : : : :
5186 TTCTCCCTAAACATGCAAGTCCAGGACACCGTTCAAGATCACAGATCTGAACAGTCC 5245
Qy 1059 ThrGlyHisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyr 1078
Db : : : : :
5246 TCA-----ATGCTATTATGAATTT 5263
Qy 1079 SerSerAspArgAspArgCysGluAlaAlaAsp-----ArgGlnProTyrHis 1095
Db : : : : :
5264 GGCCAGAAATCTCTGACCAATCCCTGCTATGACTTCAGTCAGCAGTCTCCAGATCAC 5323
Qy 1096 ArgSer-----ArgSerThrGluGlnArgProLeuLeuGluArgThr 1109
Db : : : : :
5324 CCTACAGGGGTGCAGGGGTGCTTCACATCACTGAAATGGGCCA-----ACT 5371
Qy 1110 ThrThrArgSerArgSerSerGluArgProAspThrAsnLeuMetArgSerMetPro--- 1128
Db : : : : :
5372 GAAGTGGACTACAGTCTCTTCATGATGCAGGACTCCAGTTTATCACATAAGATACCACCT 5431
Qy 1129 -----SerLeuMetThrGlyArgSer 1135
Db : : : : :
5432 ATGAGAGCGCTCTACACCCAGATAATGATCTTCTGAGCTCATCTCAGTATCTCAG 5491
Qy 1136 AlaProProSerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSer 1155
Db : : : : :
5492 GTAGAGGCTCCCGTCCACCTCTCTGCTCAT-----ACCCCTTCTCAGATCGCTCT 5545
Qy 1156 ProSerSerThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLys 1175
Db : : : : :
5546 CCT----- 5548
Qy 1176 GlyThrLeuGluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeu--- 1194
Db : : : : :
5549 -----CTCCAAGAGATACTCTATCCGATGTGCTCTCCAGAGATATCTCTTATAT 5602
Qy 1195 -----AsnLysTyrLysGlnValAlaGlySerAspProArgLeuGluGln 1209
Db : : : : :
5603 GCCTCACTCACTCTGAAAAAGTCAAGTCTGGAAGGAGAGAGACTCTCTCCAAATCT 5662
Qy 1210 AspTyrHisSerLysTyrArgSerGlyTrpAspProHisArgGlyAlaAspThrValSer 1229
Db : : : : :
5663 GATATC-----TCTCCACTCACCACGAGAGTCTCTCTCTTATATATCA 5707
Qy 1230 ThrLysSerSerAspSerAspValSerValSerAlaValSerArgThrSerSerAla 1249
Db : : : : :
5708 CCTACTTTTTCAGATCT-----ACCTCTGAGTCAAGAGAGAAACAGCAACT 5755
Qy 1250 SerArgPheSerSerThrTyrMetSerValGlnSerGluArgProArgGlyAsnArg 1269
Db : : : : :
5756 TGCACAGATTCTCTCTCCCAATAGATGCAGCATCCGACAGGCCCTATGGCTTCCGT 5815
Qy 1270 LysIleSerValPheThrSerLysMetGln-----AsnArgGlnMetGly 1284
Db : : : : :
5816 ---GCCTCAGTGTATTTCGATACATGCACACCATCTAGCTTGAATAGAGATTGTCC 5872
Qy 1285 ValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMetCys----- 1301
Db : : : : :
5873 ACACCTGGCTCGAGAGAGCAGTGGAGGAGACACACCTGCTGACTTTGATGATGCCTAT 5932
Qy 1302 ---SerLeuGlnLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThr 1320
Db : : : : :
5933 CAAAAGCCTCAGGAAACACACAGCTCCCGAGATGAAGAGATTATGACTATGACTTAT 5992
Qy 1321 SerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSer 1340
Db : : : : :
5993 GAGNAGACACCCGAGCTCAGATGTGGTGGCTATTACTATGAGAGATAGAGAGAAC 6052
Qy 1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArg 1360
Db : : : : :
6053 ACAAATCTCCAGTGCAGTGGCTACTCTAT----- 6085
Qy 1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
Db : : : : :
6086 GAGACCATTTGGAAATCTACCAAGCCCTGGAAGATGGTGACTATTCTTGAATATTATT 6145
Qy 1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsn 1400
Db : : : : :
6146 GAGNAGACACCCGAGCCCTCGAAGGGTGGTACTCATATGACATAAGTGAAGAAC 6205
Qy 1401 LeuIlePheProGlyVal-----ArgLeu 1408
Db : : : : :
6206 ACCAGCCCCCGAAGTGTGTTACAGCTATGAAAAGACTGAGAGGTCTAGAGGCTT 6255
Qy 1409 AlaSerAsp-----SerGlnPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeu 1426
Db : : : : :
6266 CTGATGACATCAGCAATGGCTATGATGACTCTGAGGATGCT----- 6307
Qy 1427 ValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMetMetAsp 1446
Db : : : : :
6308 ---GGCCACACACACTTTGGGACCCCGACT-----CTTATGAA 6346
Qy 1447 LysLysGlyGlnLeuGluValGluIleAlaArgGlyLeuValValLysProGly 1466
Db : : : : :
6347 CCCTGAGAAATTCAGTTCCCTG-----AGT 6376
Qy 1467 SerLysThrLeuProAlaProTyrVal-----LysValTyrLeuLeuAspAsn 1482
Db : : : : :
6377 CTGAAGGTATTCTATGAGACATCTACAAAGACACCAACCCCTGATATCTCCACAT 6436
Qy 1483 GlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeuGluProLeu--- 1501
Db : : : : :
6437 ACTGTT-----ACGAGCTCGAGAAATTCATAGAACCCCTCAGG 6478
Qy 1502 -----TyrGlnGlnLeuLeuSerPheGluGluSer 1511
Db : : : : :
6479 CATCCACATATCTCAGACACTTCAGACTTACTGCTACATGCGAG-----AAAGAGT 6532
Qy 1512 ProGlnGlyArgValLeuGlnIleIleValTrpLysAspTyrGlyArgMetAsp----- 1529
Db : : : : :
6533 CCCCTCAGAGCCCGTCAGAGTGCATTTATGCTCTGCTCTCTTGTGTGATACAGC 6592
Qy 1530 -----HisLysSerPhe-----MetGlyValAlaGlnIleLeuLeuAspGluLeu--- 1544
Db : : : : :
6593 ACCCAAGACAGAGCTTTCCCTCTTTCATTAATCCCAATCTCTCTTGAAGTGGTTGCCA 6652
Qy 1544 ----- 1544
```

Db 6653 GTGAAGAACCCACTGAAGAACTGAAAGCCCTCACTCAATCAGGGGAGCCCCACCGC 6712
Qy 1545 -----GluLeuSerAsnMetValleGlyTrpPheLysLeuPheProProSerSerLeu 1562
Db 6713 CTCAGGAGGAAAGCAACAGGGCCGACAGTGTGATGAAGAAACCCCTCCACCTCAGTCAGCG 6772
Qy 1563 VallapPro-----ThrSerAlaProLeu 1570
Db 6773 AGTCAGCCCCATCCAGACCCGACTGTGATGTTCCCGGAGACTGAAGAGTGCCCTCCA 6832
Qy 1571 ThrArgAlaSerGlnSerSerLeuGluSerThrGlyProSer 1586
Db 6833 TCACGGCCGATGCCAATATCGACTCTGAAGACGAGTCGAACCACTCC 6880

RESULT 8
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Alignment Scores:
Pred. No.: 1.44e-08 Length: 9025608
Score: 295.50 Matches: 422
Percent Similarity: 32.86% Conservative: 206
Best Local Similarity: 22.08% Mismatches: 712
Query Match: 3.57% Indels: 580
DB: 14 Gaps: 92

US-09-617-099B-1 (1-1590) x US-10-156-761-1 (1-9025608)
Qy 4 ProLeuGlyProArgGlyArgProAlaPro-----ThrProAla-Al 17
Db 3803595 CCATCGGGTCTCTGGAAGACGATCTGCACCCGCGGCGCGACGCGCTTCGACTCGGCTTTGC 3803654
Qy 17 aSerGlnProProGlnProGluMetProAspLeuSerHisLeu----- 32
Db 3803655 GCMGGCTCCGAGCCGCTGCCGCAACAGTTGACCGACCGCCCTCGGAGCGGCGAGTT 3803714
Qy 33 -----ThrGluGluGluArgLysIleLeuAlaValMetAspArgGlnLysLys-- 49
Db 3803715 CGAGCAGCTCGAAAGGGTCTGCTGACTTCCCGACCCCGGACTCCCGCCACGAGCCGAGG 3803774
Qy 50 -----GluGluGluGluGlnSerValLeuLysIleLysGluGluHis----- 64
Db 3803775 TCTCCCGCGCGCGAGTGTGAGGTCCACGCGCGTGCAGCGGTACACACTCCCGCCCGCC 3803834
Qy 65 -----LysAlaGlnProThrGln 70

Db 3803835 GCTTGAACGCGCTCCCTTTTGAGAGGGGAAAGGTCTTGGTGAAGTCCGCGACCCCTCAACA 3803894
Qy 70 nTrpPhe-----ProPheSerGlyIleThrGluLeuValAsnAenValLeuGlnPr 87
Db 3803895 CTGTTTGGTCTGCGGGAGTTGCGCGCGCGCGCGCTGCTGCTGAGACGCGGTGACGGGCC 3803954
Qy 87 oGlnGln-----LysGlnProAsnGluLys--GluProGlnThrLys 100
Db 3803955 CGAGCAGCTCCGCGTGTACGGCTCCCGCGCGCGCTGCGCAACACCTCAGCGCGCTTC 3804014
Qy 100 sLeu-----HisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluse 119
Db 3804015 ACTCGACCACTCCCTCCCTCGTACATGACCGCACCGCGCTCGGCATCCCGGCGATCACC 3804074
Qy 119 rGln-----GlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHi 136
Db 3804075 CCAGTGTGTGGCTGACCAAGCAGGA--CAGCGCGCGCGGTCTCC--GCTGGCGCGTGC-- 3804128
Qy 136 sLysThrLysPheAlaAspGlyCysGlyHisAenCysSerTyrCysGlnThrLysPheCy 156
Db 3804129 -GCAGCAGCTCGAGCACTCGGCGCTGTGATGG-----TGACGTCCAGGGCGGTGTC 3804178
Qy 156 sAlaArgCysGlyGlyArgValSerLeuArgSerAsnLysVal-----MetTrpValCy 174
Db 3804179 GGCTGCTCGCGAGGATCACGTCCGGCTGCTTCGCCATCGCATGGCGATCATGGCGGCG 3804238
Qy 174 sAsnLeuCysArgLysGlnGlnGluLysLeuThrLysSerGlyAlaTrpPheTyrAsnSe 194
Db 3804239 TGGCGCATGCGCGCGAGA-----ACTGTCGGAAGAGCGCCGACCCGCGGAA 3804286
Qy 194 r-----GlySerAsnThrLeuGln 200
Db 3804287 GCGGGGATGCCAGCAGGTCCAGCAGTTCGCGGGCGGTTCGCGGGCGCGCTCCCGGAGC 3804346
Qy 200 nGlnProAspGlnLysVal-----ProArgGlyLeuArgAsnGluAlaProGlnGln 218
Db 3804347 ACGCTTGTGCGCGCGGTACGCGCTCGCGATCTGTCGCGCAGCGCGGTGACCGGTGTG 3804406
Qy 218 uLysLysAlaLysLeuHisGlnProGlnPheGlnGlyAlaProGlyAspLeuSer-- 237
Db 3804407 AACCGGAGAGCGGTCTCTGGAAGACCATGCGCGACGCGCTCCCGCGATCCGCGTCCG 3804466
Qy 238 -----ValProAlaValGluLysGlyArgAlaHisGlyLeuTh 250
Db 3804467 TCCCTTTCGCGCGCGCGACCGACTTCTCGCGCGTTCGACCGCTGACGACCTTGGACGGTC 3804526
Qy 250 rArgGlnAspThrIleLysAsnGlySerGlyValLys----- 262
Db 3804527 GCGTTGTGCGCGCAGCG 3804586
Qy 263 -HisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSerArgAs 282
Db 3804587 TCACCGACCGCGCGAGGACTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3804646
Qy 282 pGlnAsnArgArgTyrGluGlnSerGluGluAspTyrSerGlnTyrValProSe 302
Db 3804647 GCGCGCGCGTGAAGTCCACGT-----TC 3804670
Qy 302 rAspGlyThrMet-----ProArgSerProSerAspTyrAlaAspArgAr 317
Db 3804671 AGACGGCGCATCTTCGAGCAGCGGAGTCCACCGGATGCGCTTCTCTGCTGCGGGAGGTG 3804730
Qy 317 sSerGlnArgGluProGln-----PheTyrGln 326
Db 3804731 GGTCCAGGCGGTCCCG 3804790
Qy 326 uGluProGly-----HisLeuAsnTyrArgAsp-----Se 336
Db 3804791 GCAGCGCGCGCGCGGTAGAGAACATCCAGGGGTACGTCCACCGCGCGCGCGCGCGCGCGCG 3804850
Qy 336 rAsnArgArgGlyHisArgHisSerLysGluTyrIleValAspAspGluAspValGluSe 356
Db 3804851 CGATCAGGTGCGCGAGCGACACGCTCGCGGTCTGCACCGCGCGCGCGCGCGCGCGCGCGCG 3804909

QY 356 rArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGln-----AlaAr 372
Db 3804910 GCGGTCTCGCTCATCAGCGCCCGCGGACCGCGATGCTCGGTGATGATCAGGAGGAG 3804969
QY 372 gTyrArgSerAspProAsnLeuAlaArg-TyrProValLysProGlnProTyrGlu----- 390
Db 3804970 GCGACGTTCCGACAGCAGTCCGCGCAGATGATCGGAACCGCGCGCCATGAACCGT 3805029
QY 391 -----GluGlnMetArgIleHisAlaGluValSerArgAlaArgHisGluArg- 406
Db 3805030 GCCCGCGGACGAATTCGCGTCTCTTACGGACAGTGTATCGACCGGACCAACCGCGCG 3805089
QY 407 -----ArgHisSerAspValSerLeuAla----- 414
Db 3805090 GTGATCATCCACCGAGACGGCGAGCAGTCCCGCGAGGACGACCCAGCCCGCTGGCG 3805149
QY 415 -----AsnAlaGluLeuGluAspSerArgIleSerLeuL 426
Db 3805150 ATCCGCGGTGAGACGGCCGCGATGATGAGGAACGACGCGGAGACGACGACCA-AGTCGAC 3805208
QY 426 euArgMetAspArgProSerArgGln-----ArgSerValSerGluArgAlaAlaM 444
Db 3805209 GAAGAACATCAGCCCGCGTCTCGTCCAGCCCCCGAGTACCCCGACAGCGCGCGAGAA 3805268
QY 444 etGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGlyGlnSerSert 464
Db 3805269 TCGCGGAGCCCGTCTCGCGAACACGCTCTCGCCGATCGGTGTCCTCCACCAAGTGTCT 3805328
QY 464 yrProGlnArgThr-----SerAsnHisSerProProThrProA 477
Db 3805329 TCCCGCAGCGTCTCGCGAACACGCTCTCGCCGATCGGTGTCCTCCACCAAGTGTCT 3805388
QY 477 rArgSer-ProIlePro-----LeuAspArgProAsp 487
Db 3805389 GGCCTCGGCCCGACCGACCGACCGGTGTAGTCGATAGGTGTACGACCAACCGGGGTAC 3805448
QY 488 MetArgArgAlaAspSerLeuArgLysGlnHis-----LeuAspProSerSerAl 505
Db 3805449 GTACGGCCCGACGAGGACGACGAGGACGACGATACGAGGACGAGGATCGCGGACGCG 3805508
QY 505 aValArg-----LysThrLysArgGluLys-MetGluThrMetLeuArgAsnAspSerLeuS 524
Db 3805509 GCCCGGATTCGTACGAAGCGCGCAGGACGCGTCCCGCGCGCCGACCGGAC----- 3805561
QY 524 erSerAspGlnSerGluSerValArgProProProProArgProHisLysSerLysG 544
Db 3805562 -----GTACGCGCGCTCGACCG----- 3805579
QY 544 lyGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAlaSerThrProG 564
Db 3805580 -----TTCGACCCAC 3805589
QY 564 luTyrThrSer-----CysAspAspValGluLeuGluSer-----GluSerValS 579
Db 3805590 AGGGGTGATCGTATGATGACGCTCGCGGATCGCGGTGCGAGCGCGGTGCGACGCTCC 3805649
QY 579 erGluLysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGly----- 595
Db 3805650 GCGAGGAACCGCGACGACGACGCGTACCGCGCGAAGAGATTGACGCGGACGACCGCGAG 3805709
QY 596 -----ValLeuSerAspSerAsnThrArgSerGluArgGlnLysLysArgMetTyrT 613
Db 3805710 TTGACGCTGTTCTTCGCGATCGA----- 3805732
QY 613 yrGlyGlyHisSerLeuGluGluLysLeuLeuLysLeuLysLeuLysLeuLysLeuLysLeu 633
Db 3805733 --GGAGATGAACCACTCTCCCATCCCGTGCAC---GCCGAAGATCGTCTCCGTGAGATG 3805787
QY 633 lyValAspThrCysSerSerThrThrLeuAsnGluGluHisSerHisSerAsp---LysH 652
Db 3805788 GCACCGGTGAAGAGCGCGAGGAGCGCTACGAGAAGTACGTCGACATGGGGATCAGCGCC 3805847

QY 652 isProValThrTrpGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuA 672
Db 3805848 GTCG-----CAGCCCGTCTCAGCAGCAGCAGCCCGCGCGCGCGAGCCCTT----- 3805894
QY 672 snLysArgLeuLysAspGlySerValProArgAspSerGlyAlaMetLeuGlyLeuLysV 692
Db 3805895 -----GGCCCGCGC-----GGTCCGCGAGA 3805913
QY 692 alValGlyGlyLysMetThrGluSerGlyArgLeuCysAlaPheIleThrLysValLysL 712
Db 3805914 TAGTCGAGCCCGACACGTGCGCATGT-----G 3805943
QY 712 ysGlySerLeuAlaAspThrValGlyHisLeuArgProGlyAspGluValLeuGluTrpA 732
Db 3805944 CTCGCTGTGTAGCGCTGTACCGCGCATCGCGAAGCGGCGAGCGA----- 3805990
QY 732 snGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleIleLeuGluSerL 752
Db 3805991 -----CAGGTGGGCGAG---CAGCAGGTGTACCGCGCGTCTCTGGAGTACG 3806033
QY 752 ysProGluPro-----GlnValGluLeuValValSerArgProIleGlyAspIle----- 768
Db 3806034 GCCCGGAACCGCGCGCTCCTTCGCGCGGTCTTCTCGCCCGTGAACGTGATCAGATCCGTC 3806093
QY 769 -----ProArgIleProAspSerThr-----HisAlaGlnLeuGluS 781
Db 3806094 CCTGTCTCGGTGTACACAGATCGTCCGCTCTTGAGAGAGATCGCCAGCAGAAACACG 3806153
QY 781 erSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSerP 801
Db 3806154 GGGGTGAGAGCAGTGTGAGAGGACGAGCAGTCCGCTCGGTCTCGTGAAGCGGTACTGC 3806213
QY 801 ro---MetSerProGlyMet-----LeuArgAsp-ValProGlnPheLeuSerGly 816
Db 3806214 CGCACCGCGCTCGAGGCTCCGATCAGCACCGCGCGACCGTTCGCGACGCTTCGCGCG 3806273
QY 817 GlnLeuSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeu 836
Db 3806274 AGCAGCAGCGCGAGGTG----- 3806291
QY 837 GlyAlaLysAspLeuProSerArgGluAspGlyArgProArgAsn-----ProTyrVal 854
Db 3806292 -----ACCCGATGCGC-----CGCCGAACTCCGCGTTCGCTCGTACG 3806330
QY 855 LysIleTyrPheLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValLys 874
Db 3806331 CCGTCGATCGTCTCGCGC---AGGTCTCGCGCGACCGCATCCCGCGCGCAGCGAG 3806387
QY 875 LysThrLeuGluProLysTrpAsnGlnThrPheIleTyrSerProVal-----HisArg 892
Db 3806388 CGGCGAGCAGCGACCGCGTGTCTGTGACGCGCATCTC-GGTGAGTGTGTGATTCACCGC 3806446
QY 893 ArgGluPheArgGluArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValArgGlu 912
Db 3806447 GCTCGCGAGACGCGCGCTCGCGCCCTCGTAGTA-----CGCGCGGGTTCGAGCGC 3806500
QY 913 GluGluSerGluPheLeuGlyGluIleLeuLeuLeuLeuThrAlaLeuLeuAspAsp 932
Db 3806501 GAACGAGGCGCAGCAGATAGAGAGGCGCAGCGCGCGCGCAGCAGCAGCGTAATAGCC 3806560
QY 933 GluProHisTrpTyrLysLeuGlnThr----- 941
Db 3806561 GAGCCGCT-TGGCCAGATAGGCGAGCGCATCTCCGTCCTCCGTCCTGTCACCA 3806619
QY 942 -----HisAspValSerSerLeuProLeuProArg---ProSerProTyrLeu 956
Db 3806620 GCATGAAGAGACATCAGCAGCAGTCCAGAACCGCTCCACGACTGCGCGTAGCGGTCTTAC 3806679
QY 957 ProArgArgGlnLeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIle 976
Db 3806680 GTTCGCGCGC---ATTACGCGATGACCCACCG----- 3806706
QY 977 SerAspSerGluValSerAspTyrAspCysGluAspGlyValGlyValSerAspTyr 996

Db 3806706 ----- 3806706
QY 997 ArgHisAsnGlyArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSer 1016
Db 3806706 ----- 3806706
QY 1017 SerAsnHisCysSerProSerGly-----SerProHisArgValAspValIleGly 1033
Db 3806707 -----GCTCCGCGCGCCACCGAAGAGACACCCACAAA-----GGA 3806742
QY 1034 ArgThrArgSerTrpSerProSerAlaProProGlnArgAsnValGluGlnGlyHis 1053
Db 3806743 CACCGACATGCCGAGAACCCCGCGCGACGCCCTGGTGGCGCG-----CAC 3806787
QY 1054 ArgGlyThrArgAlaThrGlyHisTyrAsnThrIleSerArgMetAspArgHisArgVal 1073
Db 3806788 CGCGGC-----CACCGCGTT 3806802
QY 1074 MetAspAspHisTyrSerSerAspArgAspArgAspCysGluAlaAspArgGlnPro 1093
Db 3806803 GCTGTGACCGGCTCGGTTCTCGACGACCAAGCGCGACGAGCGCGCAGA-AGAGCGCG 3806861
QY 1094 TyrHisArgSerArgSerThrGluGlnArgProLeuLeuGluArgThrThr-----ThrArg 1112
Db 3806862 -----CCCCGCGCGGAGCCGACGACGATCAACGCCACCCGCG 3806897
QY 1113 SerArgSerSerGluArgProAspThrAsnLeuMetArgSerMetProSerLeuMetThr 1132
Db 3806898 TCAGCGACATCCAGCAGCGCGGCACCTCAAGTTCGGGATCTCCCGATACATCGCGACT 3806957
QY 1133 GlyArgSerAlaProProSerProAlaLeuSerArgSerHisProArgThrGlySerVal 1152
Db 3806958 TCAACCGCTACACCGCCACG-----GCAACGAGGTGACGCGCGCGTCAACGACGAGTCAACGACGAGC 3807011
QY 1153 GlnThrSerProSerSerThrProGlyThr-GlyArgArgGly-----ArgGlnIle 1169
Db 3807012 TGGTCAGCCCGAGCTCTTCAACCGCGCGCGCGCGCGCGTCCGCGCGAAGAGAACT 3807071
QY 1169 uProGlnLeu-----ProProLysGlyThrLeuGluArgSerAl 1182
Db 3807072 TCCTGCTCCCGCTCGGTGACCTCCACCAGCCCGCCAGGTCTCCTACAGTCAATC 3807131
QY 1182 aMetAspIleGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaG 1202
Db 3807132 CCAAGGCCAAGTGTCCGACGCGCGCCCTGAGCTGGCGGACACTCCGTCGCCGTGGA 3807191
QY 1202 ySerAspProArgLeuGluGln-AspTyrHisSerLysTyrArgSerGlyTrpAspPro- 1221
Db 3807192 AGGCCACGAAACGACGACAAAGAGTACGAGCGCGCCCAACACCTCCGCGTACGACCGA 3807251
QY 1222 -----HisArgGlyAlaAspThrValSerThrLys-----SerSerAspSerAspV 1237
Db 3807252 TCAGCGCGTCCGCGCGGCGACCGACCGCCACAGTCAAGTCACTTCTCTCGCGT 3807311
QY 1237 alSerAspValSerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerT 1257
Db 3807312 AGCCCGACTG-CAAAGCGGCTCTTCGACCCCACTGCTCCGCGCGCGG-GGATCGACACCC 3807369
QY 1257 yrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerL 1277
Db 3807370 GGCCTAAGTTCAACAAGGGTGGACCGAGAGATCCCG-----TCACCGG 3807414
QY 1277 yMetGlnAsnArgGlnMetGlyValSerGly-----LysAsnLeuThrLysSerThrS 1295
Db 3807415 CGCCGCTTCAAGATCTCCGCTACGACAGAGCCCGCCAGACCATCACGGGGTCCCGA 3807474
QY 1295 erIleSerGlyAspMet-----CysSerIleGluLysAsnAspGlySerGlnSerAspT 1313
Db 3807475 CCCCCAGTGTGGGGTACGAAGCCCAAGTGGACTCGTCTTCCGCGCCCTGGACTT 3807534
QY 1313 hrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgSerSerIleGlyAlaLysM 1333
Db 3807534 GACGAGTTCCGACGAGGAGC 3808561
Db 3807535 CAGCGCTGGAGCGGACGCTACCTCAACAGAGGATCGAGTACGCTCCGCGATCTCC 3807594
QY 1333 etValAlaIleValGlyLeuSerArgLys----- 1342
Db 3807595 CGAGGACTCAAGCGGCTTGCACAGCGCGCGGACGACATCGACCATCGCGCGCGCGT 3807654
QY 1343 --SerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyGlyLys-----LysLeuA 1360
Db 3807655 GGACGAGTGCATATCACC-CTGAACGCGCGCGCGCGCGCTGAAGACATCGCGGTCC 3807713
QY 1360 rgSerThrValGlnArgSerThrGlu-----ThrGlyL 1371
Db 3807714 GCCAAGCGTCCACGACGCCATCGACCGCAAGCGCATCGCCACGCTTCGGCAAGGACC 3807773
QY 1371 euAlaValGluMetArgAsnTrp-----MetThrArgGlnAlaSerArg 1386
Db 3807774 TCCCGTTCGAGTGAAGACGTAACGCGGACCTTCTTTCATGCCAACCGACCGCGTACC 3807833
QY 1386 luSerThrAspGlySerMetAsnSerTyrSerSerGlu-----GlyAsnLeuIle--P 1403
Db 3807834 AGGACAACTCCGGTACGTACGCGACGTACGACGTGAGGCGCGCGGCGAACTCTCTCGACG 3807893
QY 1403 heProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLeuGlyP 1423
Db 3807894 CGCGCGGTGGAAGACCAAGCGCGACGCGCGGTGAAGAC----- 3807934
QY 1423 roAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValG 1443
Db 3807935 -----GGCAAGCAGCTCGCCCTCACCTACGCACTCAGCGCGCGGACACCT 3807980
QY 1443 lyMetMetAspLysLysGlyGlnLeuGluValGluIleLeuArgAlaArgGlyLeuValV 1463
Db 3807981 CCGCGCAGAGACGACGCGCGGAGCTGGTGGCGGCGAGCTCGCGCGCTCGGG-----A 3808034
QY 1463 allysProGlySerLysThrLeuPro-----AlaProTyrValLysValT 1478
Db 3808035 TCAGAGTTCGATCAAGAGTGGCGACGACGAGTACTTCTTCAAGTACGTCAACATCG 3808094
QY 1478 yrLeuLeuAspAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrL 1498
Db 3808095 GCAACTTCGAC-----CTGTGCTGCTGGTGGTGGACGACAGATCTTCCGCTCCGAGG 3808148
QY 1498 euGluProLeuTyrGlnGln----- 1504
Db 3808149 CGTACCCGATCTACCGCGACCGCGCAAGAACCTGTACGAGAACTACGCTCGGTGCG 3808208
QY 1505 -----LeuLeu-----SerPheGluGluSerProG 1513
Db 3808209 GCTCCCGGAGATCGACACCTGCTGAAGAGCGGGTGAAGACCGACCGACCGCGCGG 3808268
QY 1513 lnglyArgValLeuGlnIleValTrpGlyAspTyrGlyArgMetAspHisLysSerP 1533
Db 3808269 AGGCCAAGATC-----TACAACGAGCGCGACAGAGATCT 3808304
QY 1533 he--MetGlyValAla-----GlnIleLeuLeuAspGluLeuG 1545
Db 3808305 GGGCCCTCGGCGACCTCATCCGCTTACCAGCGCGCCCGAGTCTCCGCTCCGCTCGA 3808364
QY 1545 luLeuSerAsn-----MetValIleGlyLysP 1554
Db 3808365 ACCTCGCCAACTACCGCGCGGGGCGCTCGCGGACGACGACTTTCACCAAGTTCGCTGCG 3808424
QY 1554 heLysLeuPheProPro-----SerSerLeuValAspP 1565
Db 3808425 TCAAGAAGAGTGAACCGTCCGCGACACCCAGCGGTGGGACCGGGGTTCAGCGCGCTTCAC 3808484
QY 1565 roThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGluSerSerThrGlyP 1585
Db 3808485 CAGTTCGGGACCCCGCCCTCGCGCGGGCGCGCCACGACGACGACGCGCGCGGAC 3808544
QY 1585 roSerTyrSerArgSer 1590
Db 3808545 GACGAGTTCCGACGAGGAGC 3808561

RESULT 9
US-10-133-937-37
Sequence 37, Application US/10133937
Publication No. US20030207278A1
GENERAL INFORMATION:
APPLICANT: Khan, Javed
APPLICANT: Kingner, Markus
APPLICANT: Peterson, Carsten
APPLICANT: Meltzer, Paul
TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING, DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
TITLE OF INVENTION: OTHER BIOLOGICAL STATES
FILE REFERENCE: 11613.560S01
CURRENT APPLICATION NUMBER: US/10/133,937
CURRENT FILING DATE: 2002-11-04
NUMBER OF SEQ ID NOS: 99
SOFTWARE: PatentIn version 3.1
SEQ ID NO 37
LENGTH: 9161
TYPE: DNA
ORGANISM: Homo sapiens
US-10-133-937-37

Alignment Scores:		
Pred. No.:	3,785-12	Length: 9461
Score:	295.00	Matches: 366
Percent Similarity:	31.3%	Conservative: 221
Best Local Similarity:	19.5%	Mismatches: 691
Query Match:	3.56%	Indels: 595
DB:	12	Gaps: 78

US-09-617-099B-1 (1-1590) x US-10-133-937-37 (1-9161)

Qy	1	MetSerAlaProLeuGlyProAArgGlyArgProAlaProThrProAlaAlaSerGlnPro	20
Db	1847	ATGTCAAAACCAAGCTGCCAAGGAGAAGACG-CTGAAAAAGGAAACAAAGGTAAGCCT	1905
Qy	21	ProProGlnProGluMetProAsp-----LeuSerHisLeuThrGluGluArgLys	38
Db	1906	GAGACAGAAGAGGAGGAAGAAAGAAACCCAAAGAAAGCTGGCTTAAAGAGGAGCAAA	1965
Qy	39	IleIleLeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeu	58
Db	1966	ACACCTATCAAGAGGAGGAAAACCAACCAAAAAGGAGAGGTGAAAAAGAGTCAAAAAA	2025
Qy	59	LysIleLysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThr	78
Db	2026	GAGATCAAGAAAGAGAGGAAAAAGAAACCCCAAGAAA-----	2061
Qy	79	GluLeuValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGln	98
Db	2062	GAGTTAAGAAAGAAACACGCCCAAGGAAGTCAAGAAAGGAAGTTAAGAGGAAGAGAG	2121
Qy	99	ThrLysLeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGlu	118
Db	2122	AAGGAAGTAAAAAGGAGGAAGGAACCCCAAAAAGCAAAATTAAGAGCTCCCTTAAGAC	2181
Qy	119	SerGlnGlnGlnGlnGluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThr	138
Db	2182	GCAAGAAATCATCTACTCTCTCTCTGTAAGCAAA-AAAAACAGC---TGCTTTAAACCC	2237
Qy	139	LysPheAlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCys	156
Db	2238	AAAAGTACC-----CAAGAAGGAGAGTCTGTCAAGAAAGATTTCTGTGTGCTGC	2285
Qy	157	-----AlaArgCys	159
Db	2286	CGGAAACCCAAAGGAGAGGGGAAAATAAAGTCAATTAAGAGGAAGCGCCGCGAGA	2345
Qy	160	GlyClyArgValSerLeuArgSer-----	167
Db	2346	GGGTGTGGCTGCAGCTGTGGCACTGGAGCCACCAAGAGAGTGTGTATGAGCGCAGCTGG	2405

168 AsnLysValMetTrpValCysAsnLeuCyArgLys-GlnGln----- 181
2406 AATAGCAGCCATTTCGGCTCCCAAGAACTCGAAGCTCAGAGGTCCTTATGTATCATCTCC 2465
182 -GluLeuLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGlnG 201
2466 TGAGGATCTTACCAGAAC-----TTTGAAGAGTTTAAAGGC 2501
201 nProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGluLysAl 221
2502 TGAAGAGTCGATGTAACAAGGACATCAGCCTCAGCTGGAGCTAATCGAAGACGAAGA 2561
221 aLysLeuHisGlnGlnProGlnPheGlnGlyAlaProGlyAspLeuSerVal----- 238
2562 GAAACTGAAGGAACT-----GAGCCAGTCGAAGCCCTACGTCATCCAGAA 2606
239 ----ProAlaValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrLeuLysAs 257
2607 GGAGAGAGAGTACCAAAAGGTCCTCGCGAGTCCCTCGATGAGGGAATCACTACCACTGA 2666
257 nGlySerGlyValLysHisGlnLeuAlaSerAspMet---ProSerAspArgLysArgSe 276
2667 AGGGAGGGCGAATGTGAACAGACACCTGAGGAGCTGGACCCCTCGAAGACGAGGAGT 2726
276 rProSerValSerArg-----AspGlnAsnA-ArgTyr-GlnGlnSerGluGlnArgG 294
2727 AGACGATTTGAAAAAATTGAAGATGAAGGAGCGGTTTTGAAGAATCTTCAGAGACTGG 2786
294 uAspTyr-SerGlnTyrValProSerAsp----- 303
2787 AGACTATGAAGAGAGGCGAGAACTGAGGAGGCTGAGAGCCAGAGAGGATCGGGAGGA 2846
304 -----GlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerG 319
2847 ACAGTATGTGAGCGCCTCCAGCACACGCCCACTGAG---GATGGAAAGTGGCCAA 2903
319 nArgGluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArg 339
2904 GCGCGAGGCTGATGCATATCATCAGGGGAGAAGAGGAGTCTGTGCCAGTGGGATGACCG 2963
339 qGlyHisArgHisSerLysGlnTyrIleValAspAspGluAspValGluSerArgAspG 359
2964 AGCCGAAGAAGACATGGATGAGGCCATTGAGAAGGAGAGAGCGCTGAACAATCTGAAGAGGA 3023
359 uTyr-GluArgGln-----ArgAspGluGluGluTyrGlnAlaArgTyrAr 374
3024 GGCTGATCAGGAGGACAAAGCTGAGATGCCAGAGGAGGAGGAATATGAG----- 3072
374 gSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyr-GluGluGlnMetAr 394
3073 -----CCGAAAAAATGGAAGCTGAAGACTA 3098
394 gIleHisAlaGluValSerArgAlaArgHis-----GluArgArgHisSerAs 410
3099 TGTGATGGCTGTGTGTGACAAAGGCTGCAGAGGCTGTGTGTGCCAGGAGCATGTGAT 3158
410 pValSerLeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspAr 430
3159 CTTCAACACACCAACCAAGCA-----CTAGGAGCCCGATC 3194
430 gProSerArgGlnArgSerValSerGluArgArgAlaAlaMet-----GluAs 446
3195 TCTGCGCGAAGAACCTGCTCTTCAATTCAATTCATGATGAGACTTTACCTGGAGGCTCAGAGAG 3254
446 nGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGlyGln----- 461
3255 CGAGGCGCCCGCTTCTGATGAGAGGAATGAGAGAAGACCAGCCTGAGGAATTCATGCCAC 3314
462 -SerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerProI 481
3315 CTCTGGCTACACTCACTACTATTGAGATATCC-----AGTGAGGCCAC 3359

Db 4995 GCTTCACATCACTGAAATGGGCA-----ACTGAAGTGAGCTACAGTCTCCTC 5042
Qy 1116 rGluArgProAspThrAsnLeuMetArgSerMetPro----- 1128
Db 5043 TGACATGACGAGCTCCAGTTTATCACATAGATACACCTATGAGGAGCGGTCTACAC 5102
Qy 1129 -----SerLeuMetThrGlyArgSerAlaProProSerProAlaLe 1142
Db 5103 CCAAGATAATGATCTTTCTGAGCTCATCTCAGTATCTCAGGTAGAGCCTCCCGTCCAC 5162
Qy 1142 uSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrProGlyTh 1162
Db 5163 CTCTCTGCTCAT-----ACCCCTCTCAGATCGCTTCTCCT----- 5199
Qy 1162 rGlyArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArgSerAl 1182
Db 5200 -----CTCCAGAGATAC 5213
Qy 1182 aMetAspIleGluArgAsnArgGlnMetLysLeu-----AsnLy 1196
Db 5214 TCTATCCGATGTTGCTCTCCAGAGATATGCTCTATATGCTCTCACTCACTCTGAAA 5273
Qy 1196 sTyLysGlnValAlaGlySerAspProArgLeuGluGlnAspTyHisSerLysTyAr 1216
Db 5274 AGTCAAAGTCTGGAAGGAGAGAGCTCTCTCCAAAATCTGATC----- 5319
Qy 1216 gSerGlyTrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAs 1236
Db 5320 -TCTCCACTCACCCACGAGAGCTCTCTCTATATATCACTACTTTTTCAGATCT-- 5376
Qy 1236 pValSerAspValSerAlaValSerArgThrSerSerAlaSerArgPheSerThrSe 1256
Db 5377 -----ACCTCTGCAGTCAAAGAGAGAAACAGCAACTTGCACAGTTCTCTCTCC 5426
Qy 1256 rTyMetSerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSe 1276
Db 5427 ACCAATAGATGACATCCGACAGGCCATGCTGCTCCGT---GCCTCAGTGTATTGCA 5483
Qy 1276 rLysMetGln-----AsnArgGlnMetGlyValSerGlyLysAsnLeuth 1291
Db 5484 TACAATGCACACCATCTAGCTTGAATAGAGATTGTCCACACCTGGCCCTGGAGAGGA 5543
Qy 1291 rLysSerThrSerIleSerGlyAspMetCys-----SerLeuGluLysAsnAs 1307
Db 5544 CAGTGGAGGGAAGACACCTGTGCTTACCTATGCTATGCTATCAAAAGCCTGAGAAACAC 5603
Qy 1307 pGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysArgArgSe 1327
Db 5604 CAGGTCCTCCAGATGAAGAAGATTATGACTATGAGTCTTATGAGAAGACACCCCGACCTC 5663
Qy 1327 rSerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlase 1347
Db 5664 AGATGTGGTGGCTATTACTATGAGAAGATAGAGAACCAAAATCTCCAAGTACAG 5723
Qy 1347 rGlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerTh 1367
Db 5724 TGCTACTCTCTAT-----GAGACCATTTGGGAAACTAC 5756
Qy 1367 rGluThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSe 1387
Db 5757 CAAGACCTCGAAGATGCTGACTATTCTTCTATGAAATTATTGAGAAGACACACGACCCC 5816
Qy 1387 rThrAspGlySerMetAsnSerTySerSerGluGlyAsnLeuIlePheProGlyVal-- 1406
Db 5817 TGAAGAGGGTGGGTACTCATATGACATAAGTGAAGAGACACACGACCCCGAGTGG 5876
Qy 1407 -----ArgLeuAlaSerAsp-----SerG 1413
Db 5877 TGGTTACGCTATGAAAAGACTGAGAGCTTAGAAGCTTCTGGATGATCATGCAATGG 5936
Qy 1413 nPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThr-LeuA 1433
Db 5937 CTATGATGACTCTGAGGAT-----GGTGGCCACACACTTGG 5972

Qy 1433 laThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLysGlyGlnLeuGluV 1453
Db 5973 GGACCCAGCTACT-----CTTATGAACACCAGCTGAGAAAATTACCAG 6014
Qy 1453 alGluIleIleArgAlaArgGlyLeuValLysProGlySerLysThrLeuProAlap 1473
Db 6015 TTTCCCTG-----AGTCTGAGAGTTATTTCTTATGA 6044
Qy 1473 roTyTrVal-----LysValTyLysLeuAspAsnGlyValCysIleAlaLysL 1489
Db 6045 GACATCTACAAACACACACGACCCCTGATATCTCCACATACTGTT----- 6091
Qy 1489 yLysTyThrLysValAlaArgLysThrLeuGluProLeu----- 1501
Db 6092 -----ACGAGACTGCAGAGAAATCCTAGAACCCCTCAGGCATCCACATATTTCTACGA 6146
Qy 1502 -----TyLysGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuG 1518
Db 6147 GACTTCAGACCTATGCTACACTGCAG-----AAAAGAAAGTCCCTCAGAACCCCGTCA 6200
Qy 1518 lnIleIleValTrpGlyAspTyArgMetAsp-----HisLysSerPhe- 1533
Db 6201 GGATGTGCAATTTATGCTCGTGTCTCTGTGTAATACAAGCACCCCAAGACAGAGCTTTC 6260
Qy 1534 -----MetGlyValAlaGlnIleLeuLeuAspGluLeu----- 1544
Db 6261 ACCCTCTTCAATTAATCCCAATCTCTTGAGTGGTTTGCCAGTGAAGAACCCACTGAAGA 6320
Qy 1545 -----GluLeuSerAsnM 1549
Db 6321 ATCTGAAAAGCCCTCACTCAATCAGGGGAGCCCAACCGCTCCAGAGAGAAAGCAACA 6380
Qy 1549 etValIleGlyTrpPheLysLeuPheProProSerSerLeuValAspPro----- 1565
Db 6381 GGCCCGACAGTGTGATGAATAACCCCTCCACCTCAGTCAGCAGTCCGATCCCGACGAC 6440
Qy 1566 -----ThrSerAlaProLeuThrArgArgAlaSerGlns 1577
Db 6441 CGACTCTGATGTTCCCGGAGAGCTGAAGAGTGCCTCTCCATCAGCGCGATGCCAATAT 6500
Qy 1577 exSerLeuGluSerSerThrGlyProSer 1586
Db 6501 CGACTCTGAAGAGAGTCCGAAACCATCC 6529

RESULT 10

US-10-252-157-218
; Sequence 218, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Paris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 218
; LENGTH: 11950
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 475473.1
; NAME/KEY: unsure
; LOCATION: 9422, 10229, 10231, 10233-10352
; OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-218

Qy	600	iAsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluG1	620
Db	3848	CTCT-----TCCATGGAGGA	3862
Qy	620	uAspLeuGluTrpSerGluProGlnIleLysAspSerGlyValAspThrCysSer	638
Db	3863	AGAC---AAATTCAGCAGATCTGCTTACGTGATCT---TACTGCCTCTGAAGT	3910
Qy	639	-----SerThrThrLeuAsnGluGluHisSerHisAspLysHisProValThr	656
Db	3911	GAAAGCCAGCACCACTTTGGACATCAAGATAGCATCTCAGCTGTTTCAAGTGAAAGGT	3970
Qy	656	pGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeu	676
Db	3971	CAGCCCATCGAAG-----AGCCGCTCCCTGAG	3997
Qy	676	sAspGlySerValProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGly	696
Db	3998	TCCATCTCCACCATCACCCTTAGAAAAAGACCCCTGGGTGAACGTAGTGTGAATCTCTC	4057
Qy	696	sMetThrGlnSerGlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAl	716
Db	4058	TCTGACGCCCAAT-----	4070
Qy	716	aAspThrValGlyHisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLe	736
Db	4070	-----	4070
Qy	736	uGlnGlyAlaThrPheGluGluValTyrAsnIleIleLeuGluSerLysProGluProG1	756
Db	4070	-----	4070
Qy	756	nValGluLeuValValSerArgProIleGlyAspIleProArgIleProAspSerThrHi	776
Db	4071	----GAGATTAAAGTCTCTCAGAGGCAGAGTAGCCCGGTCTCTCTGAGGTGACCCA	4126
Qy	776	sAlaGlnLeuGlnSerSerSerSerSerPheGluSerGlnLysMetAsp-----ArgPr	794
Db	4127	AGAAAGTAGTTGAAGAACATTGTGTAGTCTCTGAGGACAGACTCTGGAAGTGTGTCAAC	4186
Qy	794	oSerIleSerValThrSerProMetSerProGlyMetLeuArgAspValProGlnPheLe	814
Db	4187	ATCTCAGTCCGTGACT-----GGCAGTGTCTGTGCACACACCTTACTATCA	4231
Qy	814	uSerGlyGlnLeuSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuLeuValTh	834
Db	4232	ATCTCTACT-----GACGAGAATCCAGTCAT-----	4259
Qy	834	rIleLeuGlyAlaLysAspLeuProSerArgGluAspGlyArgProArgAsnProTrpVa	854
Db	4260	-----CTCCCTACAGAAGTCATTGAAAAACCA-----CCAGCAGT	4294
Qy	854	lLysIleTyrPhe---LeuProAspArgSerAspLysAsnLysArgArgThr-----	870
Db	4295	TCCAGTGAAGTTTGAATTCAGTCATGCCAAAGATGAGAATGAAAGGGCTTCAGTAAGCCC	4354
Qy	871	-----LysThrValLysLysThrLeuGluProLysTr	881
Db	4355	CATGGATGAGCCCGTGCCTGACTCAGAGTCTCTATTGAAAAAGTTTGTCTCTTTACG	4414
Qy	881	pAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeuG1	901
Db	4415	CAGCCGCCCTCATTTGATCCGAGTCTGCTTATGAAAGTTTT-----	4457
Qy	901	uIleThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPhe-----	917
Db	4458	-CTAAGTGTGATGACAAAGGCTTCTGCGCAGAGTGCAGAAAGTCTTTTGAAGAAAAGAG	4516
Qy	918	-----LeuGlyGluIleLeuLeuLeuGluThrAlaLe	929
Db	4517	TGGAAACACAGGCTCTCCAGACCAAGTAAGTCAGTTCTGAAATGACTTCTACTAGTCT	4576
Qy	929	uLeuAspAsp-----GluProHisTr	936

4577	DB	TTACCAAGACAAACAGGAAGGAAAGACAGAGACTTTTGCACCAATAAAGACAGACTTTGG	4636
936	QY	pTyrLysLeuGlnThrHisAspValSerSerLeuProLeuProArgProSerProTyrLe	956
4637	DB	CCAGAAAGAAAGAACTGATGATGTGAAGCGTACAGT---TCTCAACCCAGCACTGGCTCT	4693
956	QY	uProArgGlnLeuHisGlyGluSerProThrArg-	968
4694	DB	GGATCAAGGAAATATTAGAGATGTTCTCCACACAATAGATGTCAGTCAGTTTGGATC	4753
969	QY	ArgLeuGlnArgSerLysArgIleSerAspSerGluValSerAspTyrAspCysGln	987
4754	DB	TTTTAAAGAACACACTAAGATGTCATTTCTGAAGGTACTGTCTCAGACAAGTCAGCTAC	4813
987	QY	uAsp---GlyValGlyValValSerAsp---TyrArgHisAsnGlyArgAspLeuGlnSe	1005
4814	DB	TCCTGTTGATGAGGCGGTAGCAGAAGACAGTACTCTCATATGGAGGGTGTGGCCTCAGT	4873
1005	QY	rSerThrLeuSerVal-----ProGluGlnValMetSerSerAsnHisCys	1020
4874	DB	GTCCACAGCCTCACTGGCTACGAGCTCATTTCCAGAGCCA-----ACAACAGTCATGT	4927
1020	QY	sSerProSer-----GlySerProHisArgValAspVal-----	1031
4928	DB	GTCTCCATCTCTGCATGCTGAGTGGCTCCCCACATTTCCACAGAAGTAGATGACTCCCT	4987
1032	QY	-----11eGlyArgThrArgSerTyrSerProSer--	1041
4988	DB	TTCACTGTCTGTTGTGCAACACACTACCATTTCCAGGAAACAGAAATGTTCTCCATCTAA	5047
1042	QY	-----AlaProPro-----Pr	1045
5048	DB	AGAAAGATGCCAAGACCGATGTCAATTTCTCCACAGATTTCTCCCTAAAAGTCCAAA	5107
1045	QY	oGlnArgAsnValGlnGlnGlyHisArgGlyThrArgAlaThrGlyHisTyrAsnThrIle	1065
5108	DB	GTCCAGGACACCCCTTCAAGATCACAGATCTGAACAGTCTCTCA-----	5150
1065	QY	eSerArgMetAspArgHisArgValMetAspAspHisTyrSerSerAspArgAspArgAs	1085
5151	DB	-----ATGTCATATGAATTTGGCCAAAGAATCTCTCTCAGACA	5185
1085	QY	pCysGluAlaAlaAsp-----ArgGlnProTyrHisArgSer-----	1097
5186	DB	ATCCTTGTCTATGGACTTTCAGTCAGATCTCCAGATCACCTTACAGTGGTGCAGCGGT	5245
1098	QY	ArgSerThrGluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSe	1116
5246	DB	GCTTCACATCACTGAAATGGGCCA-----ACTGAAGTGGACTACAGTCCCTC	5293
1116	QY	rGluArgProAspThrAsnLeuMetArgSerMetPro-----	1128
5294	DB	TGACATGAGGACTCCAGTTTATACATAAGATACCACTATGAGGAGCGCTCTACAC	5353
1129	QY	-----SerLeuMetThrGlyArgSerAlaProSerProAlaLeu	1142
5354	DB	CCAAGATAATGACTTTCAGTCACTCAGTATCTCAGGTAGAGCGCTCCCGCTCCAC	5413
1142	QY	uSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrProGlyTh	1162
5414	DB	CTCTTCTGCTCAT-----ACCCCTTCTCAGATCGCTTCTCTCT-----	5450
1162	QY	rglyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArgSerAl	1182
5451	DB	-----CTCCAAGAAGATAC	5464
1182	QY	aMetAspIleGluGluArgAsnArgGlnMetLysLeu-----Asnly	1196
5465	DB	TCTATCCGATGTTGCTCTCCAGAGATATGCTTATATGCTCTCACTCACTCTGAAAA	5524
1196	QY	sTyrLysGlnValAlaGlySerAspProArgLeuGlnAspTyrHiserLysTyrAr	1216

Db 5525 AGTCAAAAGTCTGGAAGAGAGAGCTCTCTCCAAATCTGATATC----- 5570
 Qy 1216 gSerGlyTrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAs 1236
 Db 5571 -TCTCCACTACCCACGAGAGTCTCTCTTTATATTCACCTATCTTTTTCAGATTCT-- 5627
 Qy 1236 pValSerAspValSerAlaValSerArgThrSerSerAlaSerArgPheSerThrSe 1256
 Db 5628 -----ACCTCTGACGTCAAGAGAGAAACAGCACTTGGCCACAGTTCCTCTTCTCC 5677
 Qy 1256 rTyrMetSerValGlnSerGluArgProArgGlyAsnArgLysLysSerValPheThrSe 1276
 Db 5678 ACCAATAGATGACATCCGAGAGCCCTATGGCTTCGGT---GCCTCAGTGTATTTCGA 5734
 Qy 1276 rLysMetGln-----AsnArgGlnMetGlyValSerGlyLysAsnLeuTh 1291
 Db 5735 TACAATGCAACACCATCTAGCTTGAATAGAGATTGTCCACACCTGGCTGGAGAGGA 5794
 Qy 1291 rLysSerThrSerLysSerGlyAspMetCys-----SerLeuGlyLysAsnAs 1307
 Db 5795 CAGTGGAGGAGACACCTGGTGACTTAGCTATGCTATCAAAAGCCTGAGAAACAC 5854
 Qy 1307 pGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgSe 1327
 Db 5855 CAGGTCCCCAGATGAGAGAGATTATGACTATGACTTCTTATGAGAGACACCCCGACCTC 5914
 Qy 1327 rSerLleGlyAlaLysMetValAlaLleValGlyLeuSerArgLysSerArgSerAlase 1347
 Db 5915 AGATGGGGTGGCTATTACTATGAGAGATGAGAGAACCAACAAATCTCCAAATGTACAG 5974
 Qy 1347 rGlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerTh 1367
 Db 5975 TGGTACTCTCTAT-----GAGACCATTTGGAAACTAC 6007
 Qy 1367 rGluThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluse 1387
 Db 6008 CAAGACCCCTGAAGATGGTACTATCTCTATGAAATATTGAGAGACACACCGACCCC 6067
 Qy 1387 rThrAspGlySerMetAsnSerTyrSerSerGlyLysLeuLlePheProGlyVal-- 1406
 Db 6068 TGAAGAGGGTGGTACTATATGATAGTGAAGAGACACACCGACCCCTCCGAAAGTAG 6127
 Qy 1407 -----ArgLeuAlaSerAsp-----SerG1 1413
 Db 6128 TGGTTACAGCTATCAAAAGACTGAGAGCTTAGAGGCTTCTGGATGACATCAGCATGG 6187
 Qy 1413 nPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThr-LeuA 1433
 Db 6188 CTATGATGACTCTGAGGAT-----GGTGGCCACACACTTGG 6223
 Qy 1433 laThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLysGlyGlnLeuGluV 1453
 Db 6224 GGACCCAGCTACT-----CTTATGAAACCACTGAGAAATATTACCAG 6265
 Qy 1453 alGluIleAlaArgAlaArgGlyLeuValLysProGlySerLysThrLeuProAlaP 1473
 Db 6266 TTTCCCTG-----AGTCTGAGGTATTTCCTATGA 6295
 Qy 1473 roTyrVal-----LysValTyrLeuLeuAspAsnGlyValCysIleAlaLysL 1489
 Db 6296 GACATCTCAAAAGACACACGACCCCTGATACTTCCACATCTGTI----- 6342
 Qy 1489 yLysThrLysValAlaArgLysThrLeuGluProLeu----- 1501
 Db 6343 -----ACGAGACTGACAGAAATCACTAGAACCCCTCAGGCATCCACATATTCCTACGA 6397
 Qy 1502 -----TyrGlnGlnLeuSerPheGluGluSerProGlnGlyArgValLeuG 1518
 Db 6398 GACTTGACCTATGCTACACTGCG-----AAAAGAGTCCCTTCAGAACCCCGCTCA 6451
 Qy 1518 lnIleIleValTrpGlyAspTyrGlyArgMetAsp-----HisLysSerPhe- 1533
 Db 6452 GGATGTGATTTATGCTCTGTCCTCTTGTGAATACAGACACCCCAAGACAGACTTTC 6511

Qy 1534 -----MetGlyValAlaGlnIleLeuLeuAspGluLeu----- 1544
 Db 6512 ACCCTCTTTTATTAATCCCAATCTCTTGTAGTGGTTTGCAGTGAAGAACCCACTGAAGA 6571
 Qy 1545 -----GluLeuSerAsnM 1549
 Db 6572 ATCTGAAAGCCCTCACTCAATCAGGGGAGCCCCACCGCTCCAGAGGAGAAACACACA 6631
 Qy 1549 etValIleGlyTrpPheLysLeuPheProSerSerLysLeuValAspPro----- 1565
 Db 6632 GGGCCGACAGTGTGATGAAGAACCCCTCCACCTCAGTCAGGAGTCAGCCCATCCAGAC 6691
 Qy 1566 -----ThrSerAlaProLeuThrArgArgAlaSerGlnS 1577
 Db 6692 CGACTCTGATGTTCCTCCCGAGACTGAGAGTGCCTCTCATCACGGCGGATGCCAATAT 6751
 Qy 1577 erSerLeuGluSerThrGlyProSer 1586
 Db 6752 CGACTCTGAGACGAGTCGGAACCATCC 6780
 RESULT 11
 US-10-132-134-1
 ; Sequence 1, Application US/10132134
 ; Publication No. US20030171562A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Farnet, Chris
 ; APPLICANT: Yang, Xianshu
 ; APPLICANT: Staffa, Alfredo
 ; APPLICANT: Zazopoulos, Emmanuel
 ; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES
 ; FILE REFERENCE: 3012-2US
 ; CURRENT APPLICATION NUMBER: US/10/132,134
 ; CURRENT FILING DATE: 2002-04-26
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 52101
 ; TYPE: DNA
 ; ORGANISM: Streptomyces platensis subsp. rosaceus
 US-10-132-134-1
 Alignment Scores:
 Pred. No.: 9,6e-11 Length: 52101
 Score: 288.50 Matches: 432
 Percent Similarity: 31.41% Conservative: 201
 Best Local Similarity: 21.44% Mismatches: 697
 Query Match: 3.48% Indels: 688
 DB: 12 Gaps: 94
 US-09-617-099B-1 (1-1590) x US-10-132-134-1 (1-52101)
 Qy 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
 Db 7891 CTGGCTGCACCATG---GTACGTGCGCGCAACGCCCGCGCGCTCGCTCGGGGT 7947
 Qy 21 ProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysIleIle 40
 Db 7948 CCCCCTGTACTGTCTGCGACCG-----CGCAGTGGCGCCCTGTGCGC 7989
 Qy 41 LeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeuLysIle 60
 Db 7990 CTAGGA-----CGGAGAACGCGCCATGAAGAGACGAGAACGCGCTCTCGCCGACGA 8040
 Qy 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
 Db 8041 CCGGGACATCGCGCTCATCGGCTGTC-----CCTGCGGTTCGCCGCTCGCGCAC 8091
 Qy 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnPro-----AsnGluLysGluProGln 98
 Db 8092 GCCCGAGGAGTCTTGAGACCACTGCGCCGAGGCGCTCGCTCATCAGCGAGTCCCGGA 8151
 Qy 99 ThrLysLeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGlu 118

[illegible]

```
QY 706 heilethrysvallyslyserleuAlaaspThrValGlyHisLeuArgProGlyA 756
Db 2120 -----TCAGATTCAATCAAGCAAAAGTAGATA 2150
QY 726 spGluValLeuGluTrpAsnGlyArgLeuGlnGlyAlaThrPheGluGluValTyrA 746
Db 2151 ATCAGCCAGAGAA-----TTAGTGCTAGTGTCTGAAGATGATGAG-----2191
QY 746 snilelleLeuGluSerLysProGlu-----ProGlnValGluLeuValValSerArgP 764
Db 2192 -----AAACCAAGTCAGAGCCAGTTACAAATGAATGCTGATCCACAGA-2233
QY 764 roiledGlyAspIleProAspSerThrHisAla-----GlnLeuGluS 781
Db 2234 --ATTTCACAGTGCTACACAACTGATAATCCATTCTTCACCTCGCAAACTCAAA 2291
QY 781 erSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSerP 801
Db 2292 GGATGAGCAAGTCTGTTCCAGCATTTCTCCAGATGAGGTGAGTGGCAGTGTGATGATG 2351
QY 801 roMetSerPro-----GlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuS 819
Db 2352 TTTATAGTGAGACTTTGGCAATCTGAA-----GTTAAAGGAATATTC 2396
QY 819 exileLysLeuTrpPheAspLysValGlyHisGlnLeuValThrLeuGlyAlaL 839
Db 2397 AGTTTCGAATTGAATATGAGTCACTAGAGAGTTGATGTTTTTGGCCCACTGTA 2456
QY 839 yAspLeuProSerArgGluAspGlyArgProArg---AanProTyrValLysIleTyrP 858
Db 2457 AGGACTTAGCAGCAGCGGATGTAATAAACAACAGCGTTTCAGACCCATATGTAAGGCCTATT 2516
QY 858 heLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuG 878
Db 2517 TGTACACAGCAAAAGCAAAATGGGCAAGAAGAAACACTCGTAGTGAAGAAACCTTGA 2576
QY 878 luProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgL 898
Db 2577 ATCTCTGTATACGAAATACCTCGGTATAAA---ATTGAATAACAAATCTTAAGACAC 2633
QY 898 rgMetLeuGluThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheL 918
Db 2634 AGAAATTTGAACCTGCTCATTTGGCAT-----CGGATACATTAAAGCCATGTTCC 2687
QY 918 euGlyGluLeuLeuLeuGluLeuGluThrAlaLeuLeuAspAspGluProHis-----935
Db 2688 TAGGGAGGTGGAACCTTGATTTGGAAACATGGGACTGGGATAACCAACAGATAAACAA 2747
QY 936 -----TrpTyrLysLeuGlnThrHisAspValSerSerLeuProLeuArg-----951
Db 2748 TGAGATGTTACCTCTGAAGCG- GAAGACAGCAGCAGTGGCCCTTGAAGCAAAACAGA 2806
QY 952 -----ProSerProTyrLeuProArg 958
Db 2807 GGTGAAATGAACACTAGCTCTCCAGTATGTCACAGA 2841
```

RESULT 13

```
US-09-814-353-19153
; Sequence 19153, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
```

```
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19153
; LENGTH: 11295
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11276, 11277, 11278, 11279, 11280, 11281, 11282, 11283, 11284,
; LOCATION: 11285, 11286, 11287, 11288, 11289, 11290, 11291, 11292,
; LOCATION: 11293, 11294, 11295
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19153
```

Alignment Scores:

Pred. No.:	6,65e-11	Length:	11295
Score:	280.00	Matches:	374
Percent Similarity:	32.84%	Conservative:	229
Best Local Similarity:	20.37%	Mismatches:	671
Query Match:	3.38%	Indels:	564
DB:	12	Gaps:	85

US-09-617-099B-1 (1-1590) x US-09-814-353-19153 (1-11295)

QY	31	HisLeuThrGluGluGluArgLysIleLeuAlaValMetAspArgGlnLysLysGlu	50
Db	364	CATGGAATCTCAGAGAGAGGTATGTCG-----GTAATCTTCGGCAGCAAGAGGAC	417
QY	51	GluGluLysGluGlnSerValLeuLysIleLys-----GluGluHisLysAlaGlnPro	68
Db	418	CAAGAAAAAGCCCTGACTGCATCAAAAGGAAAACTTTCTTTGGCATGCAGATTGAAGTA	477
QY	69	ThrGlnTrpPheProPheSerGlyIleThrGluLeuValAsnValLeuGlnProGln	88
Db	478	ACAGCATGGATA-----GGTCCAGAAAACAGAAAGTGAATAATCAATTTCCGCCCTTG	528
QY	89	GlnLysGlnProAsnGluLysGluProGlnThr-----	99
Db	529	GATGAAGGATAGATGAATTTCAACCCCAAGCAACAACTCTCTTTATGGCAACCTT	588
QY	100	-----LysLeuHisGlnGlnPhe-----GluMetTyr	108
Db	589	GAATAAACCACTACTTACCATGACCTTCGCAACATCTTCCAGCGCTTTGGAGAAATTGTG	648
QY	109	LysGluGlnValLysLysMetGlyGluGlnSerGlnGlnGlnGlnGlnGlnGlnGlnGln	128
Db	649	GATATTGACATTAAAGAAAGTAAATCGAGTTCCTCAGTATGGCTTTCTGCAATATCTGTAT	708
QY	129	AlaProThrCysGlyIleCysHisLysThrLysPheAlaAspGly-----CysGlyHis	146
Db	709	-----ATTGCTAGCGTTTGTAAAGCTATTAAAGATGGATGGGGAATATCTTGGAAAT	762
QY	147	AsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGlyGlyArgValSerLeuArg	166
Db	763	AAT-----CGCTCAAGCTGGGT	780
QY	167	SerAsnLysValMetTrpValCysAsnLeuCysArgLysGlnGlnGlnGlnGlnGlnGln	186
Db	781	TTTGGAAAGAGCATG-----CCTACA	801
QY	187	SerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGlnGln-----	201
Db	802	AACTGGGTGGCTAGATGGGCTTTCTTCGAATGTGTGATCATGATTTTAAACACGACAT	861

Qy	202	----	ProAspGlnLysValPro	-----	ArgGlyLeuArgAsn	212
Db	862	TTCTGCCGATATGGCCCTGTGGTAAAGGTGGTGTGTACCGCTTTAAAGGCATG	-----	ArgGlyLeuArgAsn	212	
Qy	213	GluGluAlaProGlnGluLysAlaLysLeuHisGlnGlnProGlnPheGlnGlyAla	232			
Db	916	-----GCCCTGGTCTCTACCAATGAAATGAATGAATGACCAAGCA	954			
Qy	233	ProGlyAspLeuSerValProAlaValGlu	-----	LysGlyArgAlaHisGly	248	
Db	955	-----GCTGTAAAGAGACCAAGGGAGGAAAAATCGGTGGGAAT	993			
Qy	249	-----LeuThrArgGlnAspThrIle	-----	ArgSerProSer	278	
Db	994	AAAATTAAAGTGGATTTTGCAAATCGGAAAGTCACTGGCTTTTATCACTGCATGGAG	1053			
Qy	256	LysAsnGlySerGlyValLys	-----	HisGlnIleAlaSerAspMetProSerAspArg	273	
Db	1054	AAATCTGGTCAAGACATCAGAGACTTTATGAATGTAGCCGAAAGAGAGGAAACGA	1113			
Qy	274	Lys	-----	ArgSerProSer	278	
Db	1114	AGGCGATCTTACGACTATAACCAAGATCGTACATATTATGAGAGTGTTCGAACTCCAGGC	1173			
Qy	279	ValSerAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGln	298			
Db	1174	ACTTATCTGTGAGTATCCAGCGGAGTATCCAGCTCGAGGAGAGAGTTTTATTTCAGAA	1233			
Qy	299	TyrValProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSer	318			
Db	1234	TGGGAACATTACCAAGGA	-----	GACTACTATGATCAACGA	1269	
Qy	319	GlnArgGluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArg	338			
Db	1270	-----TACTACGATCACTCT	-----	CGGGAATACACGGATTACAGG	1305	
Qy	339	ArgGlyHisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAsp	358			
Db	1306	-----AATGAT	1311			
Qy	359	GluTyrGluArgGlnArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsn	378			
Db	1312	CTTTATGAACAAGATATTAGGATATAGTTACAGGCAAGGGNACGA	-----	1359		
Qy	379	LeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGlu	398			
Db	1360	-----GAAAGAGAACCTGAAAGATTTGAG	1383			
Qy	399	ValSerArgAlaArg	-----	HisGluArgArgHisSerAspValSerLeuAlaAsnAlaGlu	417	
Db	1384	TCTGACCGGACAGAGACCATGAGAGAGG	-----	CGAATTGAA	1422	
Qy	418	LeuGluAspSerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerVal	437			
Db	1423	CGAAGTCAAAGTCTCTGTTCATTCCGACGCTCCACAGAGTCTGTGGAGCGCTCTCCCTCTCAG	1482			
Qy	438	SerGluArgArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGlu	457			
Db	1483	GCAGAGAGGTTGCCAGTGTATTCGAGGAGCGCTTTACAGC	-----	1524		
Qy	458	AlaGlnGlyGlnSerSerTyrProGluArgThrSerAsnHisSerProProThrProArg	477			
Db	1525	-----CGATCCTCAGACCGGAGTGAAGCTGTAGTCTCA	1557			
Qy	478	ArgSerProIleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGln	497			
Db	1558	CTCTCCCTCCCAAGATATGAGAAATCGACAAGTCTGTTGGAGCGCTATACAAAAAT	1617			
Qy	498	HisHisLeuAspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMet	517			

QY	518	LeuArgAsnAspSerLeuSerAspGlnSerGluSerValArgProProArg	537
DB	1678	ATACGGAAGAAAAGTGGAAAAGGACAAAACCTCAC	1713
QY	538	ProHisLysSerLysGlyGlyLysMetArgGlnValSerLeuSerSer	555
DB	1714	-----AAGCGAAACCCAAAGGAAGGTTCACTCCCTTAGTTCTTCAGTCTTCAGAAACG	1767
QY	556	-----GluGluGluLeuAlaSerThrProGluTyr-----ThrSerCysAspAspValGlu	572
DB	1768	GACCAAGAAATGAGCGAGAGCAAGCCCTGAAAGCCAGAGATGTGAAT-----AAA	1821
QY	573	LeuGluSerGluSerValSerGluLysGlyAspSerGlnLysGly	587
DB	1822	CTGACGACAGAGAAAGCTGACAAAGAGGGAATAGCGAAMAAACCCCTGGAACTCATGCCCT	1881
QY	588	-----LysArgLysThrSerGluGluGlyValLeuSerAspSerAsnThrArg	603
DB	1882	TGCGTGGTTTGACTCGAGTGAAGAAAGAGGAAAGGTCAATTGAC-----CACACTCCCT	1938
QY	604	SerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGluAspLeuGlu	623
DB	1939	GTGCAAAAGTGAAGCCCAAGCTT-----	1962
QY	624	TrpSerGluProGlnLysAspSerGlyValAspThrCysSerSerThrThrLeuAsn	643
DB	1963	-----GATAAT-----CACACTGTCAAATCTCTCGCCCTGGAC	1995
QY	644	GluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLysAspGlyAsp	663
DB	1996	CAGAAACTCTCAGGCTCTCTCAGACGGAGCCCTGCAAAATCTCAGTGTCTTAAA-----	2046
QY	664	ArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAsp	683
DB	2047	-----CTGGAAATCAGTTAGATGAAA-----GTACCAAGGAA	2079
QY	684	SerGlyAlaMetLeuGlyLeuLysValValGlyLysMetThrGluSerGlyArgLeu	703
DB	2080	AAGGGGCTTTCAAGCCATGTTGAAGTGGTG-----GAGAAGGAAGCAGGCTT	2127
QY	704	CysAlaPheIleThrLysValLysGlySerLeuAlaAspThrValGlyHisLeuArg	723
DB	2128	-----AAAGCCAGGAAG-----CACCTCAAG	2148
QY	724	ProGlyAsp-----GluValLeuGluTyrPAsnGly	733
DB	2149	CCTGACGAGCTGCAGATGGGTGAAGTCTGTGGATCTGGAAGCTGGAGCCAGGAAA	2208
QY	734	ArgLeuLeuGlnGlyAlaThrPheGlu-----	742
DB	2209	AGGCGCTTTGCAGATTCCAATTTAAAGACAGAAAACGCAAGGTCGAAGAAAAGC	2268
QY	743	-----GluValTyrAsnIleLeuLeuSerLysProGluProGlnVal-----	757
DB	2269	AGTCAGAGATGGAGGATGTCGCGTCTTTCAAAGAACGAGCTGACGTCTCTTAGA	2328
QY	758	GluLeuValValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAla	777
DB	2329	GAGTCAATCTGCTGAGGGAA-----GGAGAGGCTCAAGAAAGCCT-----GTG	2373
QY	778	GlnLeuGluSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSer	797
DB	2374	AGGAAAGAAATCTTAAAGAGAAATCTAAAAAATCAAACTGGACAGA-----	2421
QY	798	ValThrSerProMetSer-ProGlyMetLeuArgAspValProGlnPheLeuSerGly--	816
DB	2422	CTTAATACTGTGCGACCCCAAGACTGTCAAGAGCTTGCCAGAGCTTGCCAGTATTTCTGTGGTCT	2481
QY	817	-----GlnLeuSerIleLysLeuThr-----	823
DB	2482	GGCTCAAGGCCAGCTCAGACCTTACAGCAAGACTGGGAGAACCAAGAGTGAATCTGTG	2541
QY	823	pPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyValAlaValAspLeuProSe	843

QY 249 -----LeuThrArgGlnAspThrIle----- 255
 Db 1945 AAAATTAAGGTGATTTTGCAAATCGGAAAGTCAGCTGCGCTTTTATCACTCATCGTAGG 2004
 QY 256 LysAsnGlySerGlyValIys-----HisGlnIleAlaSerAspMetProSerAspArg 273
 Db 2005 AAATCTGGTCAAGACATCAGAGACTTTTATGAATGTAGCCGAAAGAGAGAGAACGA 2064
 QY 274 Lys-----ArgSerProSer 278
 Db 2065 AGGGCATCTACGACTATTAACCAAGATCGTACATATTATGAGAGTGTTCGAATCCCGC 2124
 QY 279 ValSerArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGln 298
 Db 2125 ACTTATCTGAGGATCCAGCGGAGCTATCCAGCTCGAGGAGAGAGTTTATTACGAA 2184
 QY 299 TyrValProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSer 318
 Db 2185 TGGGAAACTTACCAAGCA-----GACTACTATGAATCAGCA--- 2220
 QY 319 GlnArgGluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArg 338
 Db 2221 -----TACTACGATGATCT-----CGGGAATACAGGGAATACAGG--- 2256
 QY 339 ArgGlyHisArgHisSerLysGluTyrIleValAspGluAspValGluSerArgAsp 358
 Db 2257 -----AATGAT 2262
 QY 359 GluTyrGluArgGlnArgGluGluGluGluTyrGlnAlaArgTyrArgSerAspProAsn 378
 Db 2263 CTTTATGAACAAGATATTAGGAATATAGTTACAGCAAGGGAACGA----- 2310
 QY 379 LeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGlu 398
 Db 2311 -----GAAAGAGAAAGTGAAGATTGAG 2334
 QY 399 ValSerArgAlaArg-----HisGluArgArgHisSerAspValSerLeuAlaAsnAlaGlu 417
 Db 2335 TCTGACCGGGACAGACACATGAGAGAG-----CCGATTGAA 2373
 QY 418 LeuGluAspSerArgIleSerLeuLeuMetAspArgProSerArgGlnArgSerVal 437
 Db 2374 CGAAGTCAAAAGTCTGTTCACTTGGCAGCTCCACAGAGCTCTGGAGCGTCCCTCTCAG 2433
 QY 438 SerGluArgArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGlu 457
 Db 2434 CGAGAGAGTTCGCGAGTGATCTCGAGAGAGGCTTTACAG----- 2475
 QY 458 AlaGlnGlyGlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArg 477
 Db 2476 -----CGATCTCTACAGCGAGTGGAGCTGTAGTCA 2508
 QY 478 ArgSerProIleProLeuAspArgProAspMetArgAlaAspSerLeuArgLysGln 497
 Db 2509 CTCTCCCTCCAGATATACAGAACTGGACAAGTCTCGTTTGGAGCGCTATACAAAAAT 2568
 QY 498 HisHisLeuAspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMet 517
 Db 2569 GAAAAGACAGATAAAGACGAACCTTTTGTATCCGGAGAGAGTGGAGAGAGAGAGCGCTTA 2628
 QY 518 LeuArgAsnAspSerLeuSerSerGlnSerGluSerValArgProProProProArg 537
 Db 2629 ATACGGAGGAAAGATGGAAAGACCAAACTGAC----- 2664
 QY 538 ProHisLysSerLysLysGlyLysMetArgGlnValSerLeuSerSerSer 555
 Db 2665 -----AAGCAGAAACGCAAGAAAGGTTTCACTCCCTAGTCTCTCAGTCTTCAGAAAG 2718
 QY 556 -----GluGluGluLeuAlaSerThrProGluTyr-----ThrSerCysAspValGlu 572
 Db 2719 GACCAGAAATAGCGAGAGCAAGCCCTGAAAGCCCGAGGATGTGAAT-----AAA 2772

QY 573 LeuGluSerGluSerValSerGluLysGlyAspSerGlnLysGly----- 587
 Db 2773 CTGACGAGAGAGAAAGCTCACAAGAGGGAATAGCGAAAAACCGCTGGAACTCATGCGCT 2832
 QY 588 -----LysArgLysThrSerGluGlnGlyValLeuSerAspSerAsnThrArg 603
 Db 2833 TCGTGGTGTTCAGTTCGAGTGAAGAGAAAGAGGAAGGTCTATTGAC---CACACTCT 2889
 QY 604 SerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGluAspLeuGlu 623
 Db 2890 GTGGAAGATTTGAAAGCCCAAGCTT----- 2913
 QY 624 TrpSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThrThrLeuAsn 643
 Db 2914 -----GATAT-----GACACTGTCAATCTTCTCCCTGGAC 2946
 QY 644 GluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLysAspGlyAsp 663
 Db 2947 CAGAAACTTCAGTCTCTCAGACGCGAGCTGCAAAATCTGACTTGTCTAAA----- 2997
 QY 664 ArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAsp 683
 Db 2998 -----CTGGAATCAGTTAGATGAAA-----GTACCAAGGAA 3030
 QY 684 SerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSerGlyArgLeu 703
 Db 3031 AAGGGCTTTCAGGCATGTTGAAGTGTG-----GAGAAGAGAGCGCGCTT 3078
 QY 704 CysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGlyHisLeuArg 723
 Db 3079 -----AAGCCAGGAAG-----CACTCAG 3099
 QY 724 ProGlyAsp-----GluValLeuGluTrpAsnGly 733
 Db 3100 CTTGAGCAGCTGCGAGTGGGTAAAGTCTGTGATCTGGAGAGCTGGAGCCAGGAAA 3159
 QY 734 ArgLeuLeuGlnGlyAlaThrPheGlu----- 742
 Db 3160 AGGCGCTTTCAGATCTCAATTTAAAGCAGAAAGCAAAACAGAGGTCAAGAAAAGC 3219
 QY 743 -----GluValTyrAsnIleLeuGluSerLysProGluProGlnVal----- 757
 Db 3220 AGTCAGAGATGGAGATGCTCGCGTGTTCATAAAAGAGCGCTGAGTGTCTCTAGA 3279
 QY 758 GluLeuValValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAla 777
 Db 3280 GAGGTCAATCTGCTGAGGAA---GGAGAGCTGAAAGAAAGCCT-----GTG 3324
 QY 778 GlnLeuGluSerSerSerSerPheGluSerGlnLysMetAspArg----- 793
 Db 3325 AGGAAGAAATCTTAAAGAGAAATCTAAATAAAATCAAACTGGACAGACTTAATCTGT 3384
 QY 794 -----ProSerIleSerValThrSerProMetSerPro 804
 Db 3385 GCCAGCCCCAAAGACTGTGAGAGCTTCCAGTATTTCTGTGGTCTGGCTCAAGGCC 3444
 QY 805 GlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIleLysLeuTrpPhe 824
 Db 3445 AGC-----TCAGACCTACAAGCAAGACTGGGA---GAATACAGGT 3483
 QY 825 AspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuProSerArg 844
 Db 3484 GAATCTGGAAAAATCAA-----GAAGTCCAATCAAAAAAGCCCATTTCCCTCAAAA 3534
 QY 845 GluAspGlyArgProArgAsnProTyrValLys---IleTyrPheLeu----- 859
 Db 3535 -----CCACAGCTCAACAGCTGCAGGTATAGATGATCAAGGA 3573
 QY 860 ProAspArgSerAspLysAsnLysArgThrLysThrValLysLysThrLeuGluPro 879
 Db 3574 CCAGAGAGAGAGACGTTAGGAAAAAATTTGTCAGTCTCTGTCAGTGAACACCTGAACGT 3633
 QY 880 LysTrpAsnGlnThrPheIleTyrSerPro----- 889


```
Db 1852 GTCTCTCTCAAGCGCGCGGAGCGCGGCGACCCCGCG---CGGAGCCCTCTCGCTGACC 1908
Qy |||||
Db 351 spGluAspValGluSerArgAspGluTyrGluArgGlnArgGluGluGluTyrGlnA 371
Qy |||||
Db 1909 CTGAAGGC-----GCCGCGCGCGGTGCCACCGCGCGCGCGTGGAGTGGCGGATGTG 1962
Qy |||||
Db 371 laArgTyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluG 391
Qy |||||
Db 1963 GACGGGGCGCGCTCAACCGCGTGTCTTCGC-----CGCCACAGTGGC 2007
Qy |||||
Db 391 luGlnMetArgIleHisAlaGluValSerArgAlaArgHisGluArgHisSerAspV 411
Qy |||||
Db 2008 AACGCCAAGTCGCGGACCTCGGTTCGCCCGCACCGCCCGTACAGGTTCAGCGAC- 2066
Qy |||||
Db 411 alSerLeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspArg 431
Qy |||||
Db 2066 ----- 2066
Qy 431 roSerArgGlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTyrS 451
Db 2067 --CATCGCAAGACAGTACCGCGCGGTACCGCGCTTGGCGCTACAAAGCGGTCTGTCATG 2124
Qy 451 erMetGluArgThrArgGluAlaGlnGlyGlnSerSerTyr-----ProGlnArgT 468
Db 2125 CGCGCGCGCGGACCGCGAGCGCGCACCGCGCGTCCACCGGACCGCGCGCGCGCGCG 2184
Qy 468 hrSerAsnHisSer-----ProProThrProArgArg- 478
Db 2185 GAGACTCCACCGCTCTCGCGGACTACGCGATCGTCCACTACAAGCGCGCGCGACGCG 2244
Qy 479 -----SerProIleProLeuAspArgProAspMetArgArgAlaAspSerL 494
Db 2245 GACTACACCGGCTCGCGGTCTACGCTGGGCGGATCTGGCGGACCGCGGAGTTCGACGACC 2304
Qy 494 euArgLysGlnHisIleuAspProSerSerAlaValArgLysThrLysArgGluLysM 514
Db 2305 TGGCGCGCGGCGAGCTTCTGTGGCGGAGCGGTACGGCGCTTCTGCTTACGTCAAG 2364
Qy 514 etGluThr-----MetLeuArgAsnAspSerLeuSerSerAspGlnSerGluSerValA 532
Db 2365 CTGAAGCGCGCGGCTCACCGTGACT-TCTCGTCTATCAACGAGGCGCGACAGGA 2423
Qy 532 rgProProPro-----ProArgProHisLysSerLysLysGlyL 546
Db 2424 CGTCTCGCGCGCGCGACGATCGACGTCAACGCGCGCGAGGTCTGGTTCGAGCAGG 2483
Qy 546 ysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAlaSerThrProGluTyrT 566
Db 2484 CAAGGAGACCGTACGGACGAGCGCGGACTACCCCGCGGAGACAGACCAAGCGCGT 2543
Qy 566 hrSerCysAspAspValGluLeuGluSerGluSerValSerGluLysGlyAspSerGlnL 586
Db 2544 CATCC-----ACTACACCGCGCGCGAGCGGACCTACCGCGTGGGACTGCACGT 2594
Qy 586 ysGlyLys-ArgLysThrSerGluGlnGlyValLeuSerAspSerAsnThrArgSerGlu 605
Db 2595 CTGACCGGCGCGCGAGCC-----CACGGACTGTGTCCAAGCCCTCGAGCGCGGTGAGGAC 2651
Qy 606 ArgGlnLysLysArgMetTyrTyrGlyGly-----His-SerLeu----- 618
Db 2652 CGACGCTTATGGCGCGGTCTTCGAGGTCCGCTACCGACCGCGCGCCACGCGCTCAGCTA 2711
Qy 619 -----GluGluAspLeuGluThrSerGluProGlnIleLysAspS 632
Db 2712 CATCATCCACAGGCGGACGAGAGGACCTC-----TCCGCGCGCGGTCTGCTGCACT 2765
Qy 632 rGlyValAspThrCysSerSerThrThrLeuAsnGluGluHisSerHis----- 648
Db 2766 CACGCGCGCGGACGATGAGGTGTGCTGTGAACGCGCGGAGGAGAACCACTCTGCTGCCGA 2825
Qy 649 -----SerAspLys 651
Qy |||||
```

```
2826 GCCCGCGGCGAGCGCGCGCGCTCTCGACCTGACCACTCCAGCGCGTCTGGATCGACCG 2885
651 sHisProValThrTyrGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLe 671
2886 GAACACGCTCGCTTGGAAACGGCTCC-----GACGCGCGCGCTCCACCGAGCTGCT 2936
671 uAsnLysArg-----LeuLysAspGlySerValProArgAspSerGlyAl 686
2937 GTCTTCGCGGAGCGGTTCGATCCGCTCAAGGACGGCTCGCTCACCGACGACGACGCG 2996
686 aMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSerGlyArgLeuCysAlaph 706
2997 GTGGCTG----- 3003
706 eIleThrLysValLysLysGlySerLeuAlaAsp-----ThrValGlyHisLe 722
3004 -----CGGCTTTCGAAGACACCGCTCACCGGCGCGCAGAGCGCGCTTCGCGCACCT 3056
722 u-----ArgProGlyAspGluValLeuGlu----- 730
3057 GAAGTCGTACACCGCTCGGTTCGTCGACCCACCGACCGCGCGGTGCGCGAGGCTCT 3116
731 -----TrpAsnGlyArgLeuLeuGlnGlyAlaLath 740
3117 CGCGCGCGCGTCTGCTCGCAACGAGCGCGGACGAGCGGTGCTG---GCCGCGAC 3173
740 rPheGluGluValTyrAsnIleLeuLeuGluSerLysProGluProGlnValGluLeuVa 760
3174 GGGCGTCCAGCTCGCGCGGTGCTCGACGATCTGTACGACGCGACGAGCGGCGACCTC-- 3231
760 lValSerArgProIle-----GlyAspIleProArgIle-----ProAspSe 774
3232 -----GGGCCCAAGTTCGCGGCGCGGCGCACCTCGCGGTGTGTGGCGCGCCACCG 3284
774 rThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGlnLysMetAspArgPr 794
3285 GCAGTCGTCTGCTGAGGTTCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3344
794 oSerile---SerValThrSerProMetSer----- 803
3345 CGGCGTCTGCTGCTGCTGCTGCGCGCGCGCGCTCTCGAGGCGCGACCGTACCGGTACGTCGT 3404
804 -----ProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSe 819
3405 GAAGGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3464
819 rIleLysLeuTyrPheAspLysValGlyHisGlnLeuIleVal----- 833
3465 GTTCGCGCTTACACCGACTCC-----GAGCGCGCGCTGCTGCTGACCTGCGACGACGAGTGC 3521
834 -----ThrIleLeuGlyAlaLysAspLeuProSerArgGluAs 846
3522 CCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3578
846 pGlyArgProArgAsnProTyrValLysIleTyrPheLeuProAspArg----- 862
3579 CGCGGAGATCCAGGAGTGCACATCCGCGACTTCTCGGTGCGCGGACGAGCGCGCGCG 3638
863 -----SerAspLysAsnLysArgArgThrLysThrVa 873
3639 GAAGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3698
873 lLysLysThrLeuGluProLysTrpAsn-----GlnThrPheIleTyrSe 888
3699 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3758
888 rProValHisArgArgGluPheArgGluArgMetLeuLeuIleThrLysTrpAspGlnAl 908
3759 CACGATCGCGGAGAGAGTCCGCGGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3818
908 aArgValArgGluGluSerGluPheLeuGlyGluIleLeuIleGluLeuGluThrAl 928
3819 GCGCGACTCCGAGAGCAGCAGGAGTGCCTCACCGCGGTCTCGCGCGCGCGCGCGCGCG 3867
```

QY	928	aLeuLeuAspAspGluProHisTrpTyrLysLeuGlnThrHisaspValSerSerLeuPr	948
Db	3868	 : : : : : : -GAGCGTACAACTGG-----GGTACGACGCCCTTACCACATACAC	3905
QY	948	oLeuProProSerProTyrLeuProArgArgGlnLeuHisGlyGluSerProThrAr	968
Db	3906	: : : : : : : GGTCCCGGAGGCTCG-----TACGCCACCGACCGACCGACCGG	3947
QY	968	gArgLeuGlnArgSerLysArgLleSerAspSerGluValSerAspTyrAspCysGluAs	988
Db	3948	: : : : : : : : : GCGCACGGTCGAGTTCGCGAGGATGGTCAAGTCGCTG-----AACCAGGA	3992
QY	988	pglyValGlyValValSerAsp-----TyrArgHisAsnGlyArgAspLeuGlnInserSe	1006
Db	3993	: : : : : : : CGGCGCTCGGGTCGTATGACGTGCTCAACACACCGCGCGCGCCCGCGAGCGCGG	4052
QY	1006	rThrLeuSerValProGluGlnValMet-----SerSerAsnHisCysSerProSerG	1024
Db	4053	: : : : : : : : : CACC---TCGGTGTCTCAGCCGGATCGTCCCGGCTACTACAGCGGCTGTCTGCCAGCGG	4109
QY	1024	lySerProHisArgValaspVallle-----GlyArgThrArg-----	1036
Db	4110	GTCCGTGGCGACGAGCACCTGTTCCGCGAAACCGCCACCGAGAACCGCATGATGGCAA	4169
QY	1037	-----SerTrpSerProSerAlaProProProGlnArgAsnValGluGlnGlyHisArgG	1055
Db	4170	: : : : : : : GCTGGTCGTGACACTCCCTCTGCTCAGCTCGGCCA-----AGGA	4205
QY	1055	lyThrArgAlaThrGlyHisTyrAsnThrIleSerArgMetAspArgHisArgValMetA	1075
Db	4206	: : : : : : : GTACAAGTCGACGGCTTCGCTTCGACC-----	4234
QY	1075	spAspHisTyrSerSerAspArgAsp-----ArgAspCysGluAlaAlaAspArgG	1092
Db	4235	: : : : : : : --TCATGGGGCACACAGCCCAAGGCCAATCTCGCGTCCGGAAGCCCTCGACGCGCT	4292
QY	1092	lnProTyrHisArgSerArgSerThrGluGlnArgProLeuLeuGluArgThrThrA	1112
Db	4293	: : : : : : : CACCGTCGGAAGACCGCGCTCGACGGGAGAAGA-----TCATCTGTACCGGGGA	4343
QY	1112	rgSerArgSerSerGluArgProAspThrAsnLeuMetArgSerMetProSerLeuMetT	1132
Db	4344	: : : : : : : GGGCTGAAGACTTCGGCGAGGTTCGCCGACGACGCCCGCTTCGTGCAGGCCACCCAGAA	4403
QY	1132	hr-----GlyArgSerAlaProProSerProAlaLeuSerArgSerHisProArgThrG	1150
Db	4404	: : : : : : : CATGGCGGCACCGGCATCGCACTTCTCCG-----ACCGGCCCGCGACGCGC	4451
QY	1150	lySerValGlnThrSerProSerSer-----ThrProGlyThrGlyArgArgGlyArgGlnL	1169
Db	4452	: : : : : : : CGTACGGGAGGCGGCCCTTTCAGCCGACCCCGGTGTCAGGGCTTCGGGTTCGGGGCT	4511
QY	1169	euProGlnLeuProProLysGlyThrLeuGluArgSerAlaMetAspIleGluGluArgA	1189
Db	4512	: : : : : : : CTACACCGACCCCAACTCTCTCGGACG-----CGAACGGCACA-CGCGCCGACGACGAAGG	4564
QY	1189	snArgGlnMet-----LysLeuAsnLysT	1197
Db	4565	: : : : : : : CCGCGCTGCTGCATCTACAGGACCTCATAGGTGGGGCTCGAGCGGCAACTCTCGCGAAGT	4624
QY	1197	Yr-----LysGlnValAlaGlySerAspProArgLeuGluGlnA	1210
Db	4625	: : : : : : : ACCGCTTCCGACTCCAGCGGCAGAGAGGTTCAGGGCTCC-----GAGGTGCG	4672
QY	1210	spTyrHisSerLysTyrArgSerGlyTrpAspProHisArgGlyAlaAspThrValSerT	1230
Db	4673	: : : : : : : ACTACAACGGCAGC---GGGGCGCGCTAC-----GGCGAGCGCCCGCGCG	4714
QY	1230	hrLysSerSerAspSerAspValSerAspValSerAlaValSerArgThrSerSerAlas	1250
Db	4715	: : : : : : : ACGCCCTCGCTACGCGCAGCGCACGAC-----AAGCAGT	4750

Search completed: November 16, 2003, 20:21:24
Job time : 18945 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 09:17:21 ; Search time 185 Seconds

(without alignments)
3793.509 Million cell updates/sec

Title: US-09-617-099B-1

Perfect score: 8285

Sequence: 1 MSAPLGRGRPAAPTAASQP.....TTRASQSSLESSTGSPYSRS 1590

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DSV=xlh

-Q=/cgn2_1/USPTO.spool/US09617099/runat_14112003.191547_23707/app_query.fasta.1.1735

-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS-human40.cdi

-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09617099 -CGN 1 1 95 @runat_14112003.191547_23707 -NCPUS=6 -ICPU=3

-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	360.5	4.4	9511	1	US-08-056-200-93
2	360.5	4.4	9511	2	US-08-800-644-93
3	311	3.8	6755	3	US-08-931-999-4
c 4	292.5	3.5	9880	3	US-08-680-897-1
5	290	3.5	4483	4	US-08-961-527-363
c 6	288.5	3.5	33529	3	US-09-144-085-3
c 7	283.5	3.4	8878	1	US-08-759-444-2
8	270.5	3.3	28958	1	US-08-458-261B-6
9	270.5	3.3	28958	1	US-08-456-837-6
10	270.5	3.3	28958	1	US-08-457-342-6
11	270.5	3.3	28958	1	US-08-457-646A-6
12	270.5	3.3	28958	1	US-08-458-076A-6

13	270.5	3.3	28958	1	US-08-764-233A-4	Sequence 4, Appli
14	270.5	3.3	28958	1	US-08-457-335A-6	Sequence 6, Appli
15	270.5	3.3	28958	1	US-08-729-214-6	Sequence 6, Appli
16	270.5	3.3	28958	3	US-09-028-934-6	Sequence 6, Appli
17	270.5	3.3	49377	1	US-08-764-233A-1	Sequence 1, Appli
c 18	266.5	3.2	80161	3	US-09-036-987A-2	Sequence 2, Appli
c 19	266.5	3.2	71989	4	US-09-443-501A-2	Sequence 1, Appli
c 20	266.5	3.2	80161	3	US-09-370-700-1	Sequence 1, Appli
c 21	266.5	3.2	80161	3	US-09-603-207-1	Sequence 1, Appli
22	265.5	3.2	8532	1	US-08-452-655B-1	Sequence 1, Appli
23	265.5	3.2	8532	3	US-08-450-582-1	Sequence 1, Appli
24	264.5	3.2	15872	3	US-09-105-537-1	Sequence 1, Appli
c 25	264	3.2	77536	4	US-09-410-551B-1	Sequence 1, Appli
c 26	263	3.2	13987	2	US-08-804-227C-13	Sequence 13, Appli
c 27	263	3.2	44377	2	US-08-804-227C-7	Sequence 7, Appli
c 28	263	3.2	44377	2	US-08-804-198-1	Sequence 1, Appli
29	261.5	3.2	9606	1	US-07-741-940-1	Sequence 1, Appli
30	261.5	3.2	9606	1	US-08-289-548A-1	Sequence 1, Appli
31	261.5	3.2	9606	1	US-08-452-65A-1	Sequence 1, Appli
32	261.5	3.2	9606	2	US-08-370-235A-1	Sequence 1, Appli
33	261.5	3.2	9606	2	US-08-449-731-1	Sequence 1, Appli
c 34	260.5	3.1	68750	3	US-09-335-409-1	Sequence 1, Appli
c 35	260.5	3.1	68750	4	US-09-568-102-1	Sequence 1, Appli
c 36	260.5	3.1	68750	4	US-09-567-969-1	Sequence 1, Appli
c 37	260.5	3.1	68750	4	US-09-568-480-1	Sequence 1, Appli
c 38	260.5	3.1	68750	4	US-09-568-486-1	Sequence 1, Appli
c 39	260.5	3.1	68750	4	US-09-568-472-1	Sequence 1, Appli
c 40	260.5	3.1	68750	4	US-09-567-899-1	Sequence 1, Appli
c 41	259.5	3.1	47981	4	US-09-679-279-1	Sequence 1, Appli
c 42	258.5	3.1	32768	4	US-08-961-527-71	Sequence 71, Appli
43	254	3.1	5200	4	US-08-978-277A-3	Sequence 3, Appli
c 44	252.5	3.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c 45	250	3.0	50937	3	US-09-428-517-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-056-200-93

; Sequence 93, Application US/08056200

; Patent No. 5616500

; GENERAL INFORMATION:

; APPLICANT: Steinert, Peter M.

; APPLICANT: Lee, Seung-Chul

; APPLICANT: Kim, In-Gyu

; APPLICANT: Chung, Soo-Il

; APPLICANT: Park, Sang-Chul

; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and

; TITLE OF INVENTION: Methods of Using Same

; NUMBER OF SEQUENCES: 117

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 620 Newport Center Drive, Sixteenth Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/056,200

; FILING DATE: 30-APR-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Fedrick, Michael F.

; REGISTRATION NUMBER: 36,799

; REFERENCE/DOCKET NUMBER: NIH054.001A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (714) 760-0404

; TELEFAX: (714) 760-9502

; INFORMATION FOR SEQ ID NO: 93:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9551 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1507..1644

; FEATURE:

; NAME/KEY: intron

; LOCATION: 1645..2511

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 2512..8070

; US-08-056-200-93

Alignment Scores:

Pred. No.: 2,278-14 Length: 9551

Score: 360.50 Matches: 324

Percent Similarity: 33.02% Conservative: 242

Best Local Similarity: 18.90% Mismatches: 568

Query Match: 4.35% Indels: 586

Dbs: 1 Gaps: 63

US-09-617-099b-1 (1-1590) x US-08-056-200-93 (1-9551)

QY 19 GlnProProGlnProGlnMetProAspLeuSerHisLeuThrGluGluGlu----- 36

Db 3163 GAGCGCGCGCGCAAGAGAGCTCCAGGAG-----GAGAGAGCAGCTA 3207

QY 37 ArgLysIleIleLeuAlaValMetAspArgGlnLysLysGluGluGluLysGluInSer 56

Db 3208 CGGAAGCTGAGCGCGCAAGAGCTGAGGAGGAGCGCCAGGAGCAAGAGCAGCAGCA 3267

QY 57 ValLeuLysIleLysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGly 76

Db 3268 AGGCTGAGCGCGCGCAGCACTAAGCGCAAGCAGGAGGAG----- 3309

QY 77 IleThrGluLeuValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGlu 96

Db 3310 -----GAGAGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3342

QY 97 ProGlnThrLysLeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGly 116

Db 3343 CAGCAGGAGAGCGCGCAGCAGCAGGAGAGAGCGCGCAGCAGCAGCAGCAGCAGCAG 3402

QY 117 GluGluSerGlnGlnGln-----GluGlnLysGlyAspAlaProThrCysGly 133

Db 3403 GAGGAGAGCGCGCAGCAGCAGCTGAGCGCGCAGCAGCAGGAGGAG----- 3447

QY 134 IleCysHisLysThrLysPheAlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThr 153

Db 3447 ----- 3447

QY 154 LysPheCysAlaArgCysGlyGlyArgValSerLeuArgSerAsnLysValMetTrpVal 173

Db 3447 ----- 3447

QY 174 CysAsnLeuCysArgLysGlnGlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsn 193

Db 3448 -----AGCGCGCAGCAGCAGCTG----- 3465

QY 194 SerGlySerAsnThrLeuGlnGlnProAspGlnLysValProArgGlyLeuArgAsnGlu 213

Db 3466 -----AGCGCGCAGCAGCAGGAGGAGGAGCGCGCAGCAGCAGCAGCAGCAGCAG 3513

QY 214 GluAlaProGlnGluLysLysAlaLysLeuHisGluGlnProGlnPheGlnGlyAlaPro 233

Db 3514 CAGGAGGAGGAGGCGCGCAGCAGCAGCTGAGCGCGCAGCAGCAGCAGCAGCAGCAG 3573

QY 234 GlyAspLeuSerValProAlaValGluLysGlyArgAlaHisGlyLeuThrArgGlnAsp 253

Db 3574 -----CAGCTGAGCGCGCAGCAGCAGCAGCTGAGCGCGCGCAGCAG 3609

QY 254 ThrIleLysAsnGlySerGlyValLysHisGlnIleAlaSerAspMetProSerAspArg 273

Db 3610 CAGCTGAGCGCGCAGCAGCAGCTGAGCGCGCAGCAGCAGCTGAGCGCGCAGCAGCAGCTG 3669

QY 274 LysArgSerProSerValSerArgAspGlnAsnArgTyrGluGlnSerGluGluArg 293

Db 3670 AGGCGCGCAGCAGCAGCTGAGCGCGCAGCAGCAGCTGAGCGCGCAGCAGCAGCAGCAGCAG 3729

QY 294 ---GluAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAsp 312

Db 3730 CACGAGCAGAAAGCAGCAGCAGCAGCGCGCAGCAGCGGCTGAAGCGCGCAGCAGCAGGAG 3789

QY 313 TyrAlaAsp-----ArgArgSer 318

Db 3790 AGGCGCGATTGCTGAAGCGCGCAGGAGCAGGAGGAGGAGCAGCAGCAGCAGCAGCAGCAG 3849

QY 319 GlnArgGluProGlnPheTyrGluGluProGlyHis-----LeuAsnTyrArgAsp 335

Db 3850 CAGCAGCTGAAGCGCGCAGCAGCAGCAGGAGGAGCGCGCAACGTTGGCTCAAGCTCGAGGAG 3909

QY 336 SerAsnArgArgGlyHisArgHisSerLysGluTyrIleValAspAspGluAspValGlu 355

Db 3910 GAGGAGAGCGCGCAGCAGCAGCAGGAGCGCGCAGCAGCAACTAAGCGCGCGCAGCAAGAGGAG 3969

QY 356 SerArgAspGlu---TyrGluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArg 374

Db 3970 AGGCGCGCAGCAGCGCTGAAGCGCGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4029

QY 375 SerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMetArg 394

Db 4030 AGCGAGCAACCACTAAGACGC-----GAGCAGGAGGAGGAGGAGCTCGACAGCTGCTG 4080

QY 395 IleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSerLeuAla 414

Db 4081 AAGCGCGAGGAGGAGGAGGAGCTCGACGAGGAGGAGGAGGAGCAGCAGCAGCTG----- 4131

QY 415 AsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspArgProSerArgGln 434

Db 4132 AAGCGCGCAGCAGGAGGAGGAGCGCGATCAGCTGCTGAAGCGCGCAGGAGGAGGAGGAGGAG 4191

QY 435 ArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArg 454

Db 4192 CAGCGGCTGAAGCGCGCAGCAGCAGGAGGAGCTCGAGCAGCGCA---CTGAAGCGCGCAGGAG 4248

QY 455 ThrArgGluAlaGlnGlyLysSerTyrProGlnArgThrSerAsnHisSerProPro 474

Db 4249 GTGAGAGACTCGAGCAGGAGGAGGAGCGCGCAGCAGCGGCTGAAGCGCGCAGGAGCAGGAG 4308

QY 475 ThrProArgArgSerProIleProLeuAspArgProAspMetArgArgAlaAspSerLeu 494

Db 4309 GAAGAGAGCGCCAC-----GAGCTGCTGAAGCGCGCAGGAGCAGGAGCAG 4347

QY 495 ArgLysGlnHisHisLeuAspProSerSerAlaValArgLysThrLysArgGluLysMet 514

Db 4348 GAGGAGAGCGCCAC-----GAGCAACTGAGCGCGCAGCAGCAGCAGGAGGAGCAGCAG 4395

QY 515 GluThrMetLeuArgAsnAspSerLeuSerSerAspGlnSerGluSerValArgProPro 534

Db 4396 GAGCAGCGCGCTGAAGCGCGCAG-----GAGGAGGAGGAGGAGGAGCTCGAGCAGCGG 4443

QY 535 ProProArgProHisLysSerLysGlyGlyLysMetArgGlnValSerLeuSerSer 554

Db 4444 CTGAAGCGCGCAGCATGATGAAGAGAGAGG-----CGCAGCAGCAGCAGCTAGCTAG 4491

QY 555 SerGluGluGluLeuAla-----SerThrProGluTyrThrSerCys 568

Db 4492 GAGGAGCAGGAACAGCGCGCGCGGATTAAGAGCCGATCCCGAAGTGG----- 4542

6022	Db	CGCGACAGACAATCCCGACGAAAGATCTCGACGACTGCTGGGTGAACACGACGAGAGAGAGAT	6081
1264	Qy	-----ArgProArgGly-----AsnArgLysIleSerVa	1273
6082	Db	CGTGAGCAAGAGAGAGCGCTGCACGACGGCCACACAGGATTTCCGACGAGGAGACAG	6141
1273	Qy	lPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSe	1293
6142	Db	CTGAGCGAGAGACGAAAGGAAAGCAGGAGGCGGACAGGAAAGTCCCAAGAGGAGAAAG	6201
1293	Qy	rThr-----Se	1295
6202	Db	CAGTGTCTGAGAGAGGAGAAAGAGAGAGAGCGCTCAAGAGACAGACAGAAATTC	6261
1295	Qy	rIleSerGlyAspMetCys	1301
6262	Db	CGCAGAGGAGACAGCTGCTCCAGGAAGGGAGGAACAGCGCTGCTCCGCCAAGAGCGT	6321
1301	Qy	-----	1301
6322	Db	GACAGAAAAATTCGCGAAGAGAGAACTGCTCCATCAGGAACAAGGGAGAAAAATTCCTCGAG	6381
1302	Qy	-----SerLeuGluLysAsnAspGlySe	1309
6382	Db	GAGGAACAGCGGCTGCGCGAGGAACGGAGAGAAAAATTCCTTAAGGAGACACGACGCTG	6441
1309	Qy	rGlnSerAspThrAlaValGlyAlaLeuGlyThr-----SerGlyLysLysArgAr	1326
6442	Db	CGCTTCAGGAGCGCGAGCACTCGTCTCAGGACGGCAGACAGAAAAATTCGCCGAGGAGGAA	6501
1326	Qy	gSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAl	1346
6502	Db	CAGCAGCTGAGCGCCCAAGACGCTGCAGACAAAAATTCGTGACAGAGGACACGACGCTGCGC	6561
1346	Qy	aSerGln-----LeuSerGlnTh	1352
6562	Db	CGCCAGGAACGAGAGAGAAAAATTCCTCGAGGAGGAACAGCAGCTGCGCCAGGAGCGTCAC	6621
1352	Qy	rGluGlyGlyGlyLysLys-----	1358
6622	Db	AGAAAAATTCGCGAAGAGGAACACGCTGCTCCAGGAAGGAGAGACAGCAGCTGCACCGC	6681
1359	Qy	-----LeuArgSerThrVa	1363
6682	Db	CAAGAGCGTGACAGAAAAATTCCTGGAGGAGGAACAACAGCTGCGCCGCCAAGAGCGTGAC	6741
1363	Qy	lGlnArgSerThrGluThrGlyLeuAlaVal-----GluMetArgAsnTrpMetThrAr	1381
6742	Db	AGAAAAATTCGCCAACACGGAACCTCGCAGTCCAGAACACGACAGAGAAAAATTCCTCGAGGAG	6801
1381	Qy	gGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsnLe	1401
6802	Db	GAAACAGAGCTGCACCGCCACGACCGCAGAGAAAAATTCCTC-CAGAGAGAACAGCAGCT	6860
1401	Qy	uilePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLe	1421
6861	Db	GCGCCG-CCAGAGCGCGGGCAACAGCGCGCTCA-----	6893
1421	Qy	uGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly-AspIleG	1441
6894	Db	GGACCGGTGACAGAAAAATTCGCCAGGAGGAGAACAGCTCGCCAGGAGGAGGAGGAACAGC	6952
1441	Qy	lnValGlyMetMet-----AspLysLysGlyGlnLeuGluValGluIleIleArgAlaA	1459
6953	Db	AGCTGAGCCCGCAAGAGCGTGACAGAAAAATTCCTGTTAGAGGAACAGAAAGTGCGCCGCC	7012
1459	Qy	rgGlyLeuValLys-----ProGlySerLysTr	1469
7013	Db	A-GGAACAAGAGAGAAAAATTCATGGAGGACGAACAGCAGCTGCGCCCGCAGGAGGCGCA	7071
1469	Qy	hrLeuProAlaProTyrValLysValTyrLeuLeuLeuAspAsnGlyValCysIleAlaLysL	1489
7072	Db	CAACAGCTGCGCCA-----GGAGGACAGAAAAATTCGCGAA-----	7107

```

Qy 1489 ysiysThrLys-ValalaArgLysThrLeuGluProLeuTyrGlnGlnLeuSer--- 1507
Db 7108 GACGAACAGCTGCTCCAGGAAAGGGAAGAACAAGCAGCTGCACGCCCAAGAGCGTGACAGA 7167
Qy 1508 -----PheGluGluSerProGlnGlyArg 1515
Db 7168 AAATTCTCGAGGAGGAACCGCAGTCGCG 7197

RESULT 2
US-08-800-644-93
; Sequence 93, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; METHOD OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,644
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/056,200
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1507..1644
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1645..2511
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2512..8070
US-08-800-644-93

Alignment Scores:
Pred. No.: 2,27e-14
Score: 360.50
Percent Similarity: 33.02%
Best Local Similarity: 18.90%
Matches: 324
Conservative: 242
Mismatches: 568

```


Qy	1309	rGlnSerAspThrAlaValGlyAlaLeuGlyThr-----SerGlyIysIysIysArgAr	1326
Db	6442	CGCCTCGAGGACCGCGACCACTCGGTGAGACCGCAGCAAAATTCOCGAGGAGGAA	6501
Qy	1326	gSerSerIleGlyAlaIysMetValAlaIleValGlyLeuSerArgIysSerArgSerAl	1346
Db	6502	CAGCAGCTGAGCGCCAGAGCGGTGACAGAAATTCGTGAGAGGAAACAGCAGGTGCGC	6561
Qy	1346	aSerGln-----LeuSerGlnTh	1352
Db	6562	CGCCAGGAACGAGAGAGAAAATTCCTGGAGGAGAAACAGCAGCTGCGCCAGGAGCGTCAC	6621
Qy	1352	rGluGlyGlyGlyIysIys-----	1358
Db	6622	AGAAATTCGCGAAGAGGACAGCTCTCCAGGAAGGAGAAACAGCAGCTGCACCCG	6681
Qy	1359	-----LeuArgSerThrVa	1363
Db	6682	CAAGAGCGTGACAGAAAATTCCTGGAGGAGGAAACACAGCTGCGCCCAAGAGCGTGAC	6741
Qy	1363	lGlnArgSerThrGluThrGlyLeuAlaVal-----GluMetArgAsnTrpMetThrAr	1381
Db	6742	AGAAATTCGCGAACAGGAACTCGGCAGTCAAGAACCAAGAGAAATTCCTCGAGGAG	6801
Qy	1381	gGlnAlaSerGluSerThrAspGlySerMetAsnSerTy+SerSerGluGlyAsnLe	1401
Db	6802	GAACAGCAGCTGCACCGCCAGCAACGCGCAGAGAAAATTCCTC-CAGGAGGAACAGCAGCT	6860
Qy	1401	uIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLe	1421
Db	6861	GCGCCG-CCAGGAGCGCGGCGAACAGCGCGTCA-----	6893
Qy	1421	uGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly-AspIleG	1441
Db	6894	-GGACCGTGACAGAAAATTCGCGAGGAGGAACAGCTGCCCGAGGAGCGGAGGAACAGC	6952
Qy	1441	lValGlyMetMet-----AspIysIysGlyGlnLeuGluValGluIleIleArgAlaA	1459
Db	6953	AGCTTGACGCCCAAGAGCGGTGACAGAAAATTCGTTTAGAGGAACAGAAATTCGCGGCC	7012
Qy	1459	rgGlyLeuValValIys-----ProGlySerIysTr	1469
Db	7013	A-GGAACACAGAGAAAATTCATGAGGAGCGAACAGCAGCTGCGCCCGCAGGAGGGCCAA	7071
Qy	1469	hrLeuProAlaProTy+ValIysValTy+LeuLeuAspAsnGlyValCysIleAlaIysLe	1489
Db	7072	CAACAGCTGGCCCA-----GGAGGACAGAAAATTCGCCGAA	7107
Qy	1489	ysIysThrIys-ValAlaArgIysThrLeuGluProLeuTy+GlnGlnLeuLeuSer---	1507
Db	7108	GACGAACAGCTGCTCCAGGAAGAGGAAGAACAGCAGCTGCACCGCAAGAGAGCGTGACAGA	7167
Qy	1508	-----PheGluSerProGlnIysArg	1515
Db	7168	AAATTCCTCGAGGAGGAACCGCAGCTGCGC	7197

419 ----- 419

L49 SerTyrCysGlnThrIlysPheCysAlaArgCysGlyGlyValSerLeuArgSerAsn 168
::: ||| |
420 AACAGGACACAACCGAAA-----GGGGGGGACGAGAAAAAAGAAGACCAA 467

L69 LysValMetTrpValCysAsnLeuCysArgLysGlnGlnGluLeuThrLysSerGly 188
||| |::| |
468 -----ACGAAGAAACACGAGAAAAAACCAACCAGAAAGAAAA----- 506

L89 AlaTrpPheTyrAsnSerGlySerAsnThrLeuGlnProAspGlnLysValProArg 208
||||| |
507 -----GAAGAAACCAAGCCAGAGACCCAAGAG- GAACGC 544

L209 GlyLeuArgAsnGluGluAlaProGlnGlnLysAlaLysLeuHisGluGlnProGln 228
::: ||| |::: |
545 -----AAAACGGCCAGAAAAGAAAAAGCAAA-----AGCCAACACACAG 586

L229 PheGlnGlyAlaProGlyAspLeuSer---ValProAlaValGluLysGlyArgAla---- 246
::: ||| |
587 GA-GAAGGAAACAAAAAAGAGACGAGCAGAAACACAAAGAGAGAGGGGAGGCCACCA 645

L247 -----HisLysLeuThrArgGlnAspThrLys 256
||||| |
646 ACGAAAAAAGAAACAAAAAGAAAGACCGGAAGCGCAAGAAACACAGACCAAC 705

L257 AsnGlySerGlyValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSer 276
::: ||| |
706 AAAACGGCAAAACCGCCACACAGAAAAACACACCCAAAAAAGACGACAGAAAAA 765

L277 ProSerValSerArgAspGlnAsnArg-----ArgTyrGluGlnSerGluGluArg 293
::: ||| |
766 AAGCAAAAAAGAAAAACGAAGGAAACAGGAACCCAGCAACAAAAACACAGCGAAAAAGCAC 825

L294 GluAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAspTyr 313
::: ||| |
826 GGAATAAAACAAACAAACACACCCCCAAACAAAGAGGCACAAATAAAGGCAACACACC 885

L314 AlaAspArgArgSerGln--ArgGluProGlnPheTyrGluGluProGlyHisLeuAsn 332
::: ||| |
886 AAAAAAAGGCACGCAAGAGAGAAAAAAGGGGAAGGAAAA----- 930

L333 TyrArgAspSerAsnArgArgGlyHisArgHisSerLysGluTyrIleValAspAspGlu 352
::: ||| |
931 -----AACCCAAGGAGGACCAAAAAAGAGGAAAA----- 963

L353 AspValGluSerArgAspGluTyrGluArgGlnArgArgGluGluGluTyrGlnAlaArg 372
::: ||| |
964 ---AAAAAATACCGGAGAAACAGAGAAAAAGGAAAAACAAACAGAGAGACCAACAA 1020

L373 TyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGlu--- 391
::: ||| |
1021 CGAAACCGGACCCGNAACCAAAAA-----AAGAAAGCCAGGCCGAGGAAAGAGA 1074

L392 -----GlnMetArgIleHisAlaGluValSerArgAlaArg 403
||| |
1075 CAAGAAAAAAGCACAGAGACACAGCGGAAGGAGAGAACCCACAAAGAACAAAAAGCAGA 1134

L404 HisGluArgArgHisSerAspValSerLeuAlaAsnAlaGlu-----LeuGluAsp 420
::: ||| |
1135 GAACAGAAAAAAGCAGCGACCAACCCCACAAAGAAAGAAAAAAGAACGACAGACAGAC 1194

L421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
||| |
1195 AAAGCAACAGAAAAAAGCGGGCAAAAAAGCAGACCAAGAGACGACCAAAAGAAAGG 1254

L441 ArgAlaIleMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
||| |
1255 AAAAAGAGGAC- GAAGAAAAAACAACAGACACACCAACAAAGAACAGCAAGAAAA--- 1310

L461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
::: ||| |
1311 GAAGACCGGAAACACGACCAAAACG-----GAAGGACCGCCCAAGAAAGAGGAGGCCA 1364

Qy	481	IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisIleu	500
Db	1365	GGAGAGCAGAAAAAGGCG-----AGAACGGCCCCCAGCAACGAAAAACAAAAAACCC	1418
Qy	501	AspProSerSerAlaValArgLysThrLysArg-GluLysMetGluThrMetLeuArgAs	520
Db	1419	AAAAGGAAAGGCGCGAGGAAAAAGCAAAAGAGGAGAAACACAGCAAGAAAAACGCG	1478
Qy	520	nAspSerLeuSerSerAspGlnSerGluSerValArgProProProProProHisLys	540
Db	1479	GGAAGGAGCGCGAAGACGGCAA---AGGAGCCACGCGGAAGACACGAAAAACAAGAA	1535
Qy	540	sSerLysLysGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAl	560
Db	1536	AAAAACAGCAAAAGCGAAAGAGAAAGAAAAAGAAAAAGAAAGAGAAAGCGCAAGACAGGA	1595
Qy	560	aSer-----ThrProGluTyrThrSerCysAspAspValGluLeuGluSe	575
Db	1596	GAGCGGACCAAAAGAGAAAAAACCAAGAGAACCC-----GA	1631
Qy	575	rGluSerValSerGluLysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGl	595
Db	1632	CGAAACAAAAAGAAAGGAAGAACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1691
Qy	595	yValLeuSerAspSerAsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGl	615
Db	1692	AAAAAGCAGAGAGACAAAAAAGAAAAAAGAAAAAGAAAAAGAAAAAAGAAAAAAGAA	1751
Qy	615	yHisSerLeuGluGluAspLeuGluTyrSerGluProGlnLysLysAspSerGlyValas	635
Db	1752	ACAAACAAAAACAGCGCAAAAAAGAAAAAGCGGAGAACAAAAAAGAACACGCGGGCAAGGA	1811
Qy	635	pThrCysSerSerThrThrLeuAsnGluGluHisSerHisSerAspLys-----HisPr	653
Db	1812	CCGGACACAGCAACGAAGGAAAAAGCGGAACAAAGAGAGAGAGAAAAAGCAAGAA	1871
Qy	653	oValThrTrpGlnProSerTysAspGlyAspArgLeuIleGlyArgLileLeuLeuAsnLy	673
Db	1872	AACACAGCGCGCAACCGCAGGAGAGGAGACGAAAAAGAAAGCAAGACGAAAAAGAAACAC	1931
Qy	673	sArgLeuLysAspGlySerValProArgAspSerGlyAlaMetLeuGlyLeuLysValva	693
Db	1932	GAAGAGAAAAAGAAAAACA-----	1951
Qy	693	lGlyLysMetThrGluSerGlyArgLeuCysAlaPheIleThrLysValLysLysGl	713
Db	1950	-----	1950
Qy	713	ySerLeuAlaAspThrValGlyHisLeuArgProGlyAspGluValLeuGluTyrAsnGl	733
Db	1951	-----CGAGCCGCGAAAGAAAAACAAAGAGAGAGAGAGAGAG	1981
Qy	733	yArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleLeuLeuGluSerLysPr	753
Db	1983	AAACACAGCCCCAGGAAAAAGACA-----GGAAAAACC	2013
Qy	753	oGluProGlnValGluLeuValValSerArgProIleGlyAspIleProArgIleProAs	773
Db	2013	GGAGAAAAACA-----GAAAGCAGA-----GGGAAAGAGCC-----CCCAA	2041
Qy	773	pSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGlnLysMetAspAr	793
Db	2049	CGAGGGAAGGCAAGAAAAAAGAAAAAGAGCCACC-----	2088
Qy	793	gProSerIleSerValThrSerProMetSerProGlyMetLeuArgAspValProGlnPh	813
Db	2089	-----ACAAAGCCAAACAGAAAGCGGAGGCGGAAAAAGAA-----	2128
Qy	813	eLeuSerGlyGlnLeuSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleVa	833
Db	2122	-----CGAAAAAGAGGACACAGAAC-----	2141

Qy	833	lThrIleLeuGlyAlaIysAspLeuProSerArgGluAspGlyArgProArgAsnProTyr	853
Db	2143	-----GAIAAAGAAACACGAGGAAGCGGAAACAGAAAAGGGCGAA	2189
Qy	853	rValLysIleTyrrPheLeuProAspArgSerAspLysAsnLysArgargThrIysThrVa	873
Db	2190	GAAACAA-----AGGAAGAAGAGAAAGAAAGAACCCGGAAGGAACCGAAGCCGGA	2240
Qy	873	LlysLysThrLeuGluProLysIstTPAsnGlnThrPheIleTyrrSerProValHisArgAr	893
Db	2241	AAGCGCAACGAGCAGCAAGAAAAACGAG-----AA	2273
Qy	893	GLuPhaArgGluArgMetLeuGluIleThrLeuTrpAspGln-AlaArgValargLug	913
Db	2274	GCACCAGAGGGAGAAAAAAGAGACAGAAAAAGCAAGAAAGCAAGACGCCAAGAAAA	2333
Qy	913	LuGluSerGluPheLeuGlyGluIleLeuIleGluLeuGluThrAlaLeuLeuAspArg	933
Db	2334	CANAAACGANA-----CAACGCACAAAAAAACAACCAA	2366
Qy	933	luProHISrtPTyrrLysLeuGlnThrHisaspValSerSerLeuProLeuProArgPros	953
Db	2367	GGCCAAACAAAGAAAGAAAGAAACGAGGGAAGCAACACACACACACCCAGGCCAC	2426
Qy	953	erProTyrrLeuProArgArgGlnLeuHisGlyGluSerProThrArgArgLeuGlnArgs	973
Db	2427	AC-----AAAAAAAAACAGCACAAAC-CANAAAC	2455
Qy	973	erLysArgIlleSerAspSerGluValSerAspTyrrAspCysGluAspGlyValGlyValv	993
Db	2456	AAAGAACAACCAAGAAAGCGCGGCGAGAGAACGACACNA-----	2499
Qy	993	alSerAspTyrrArgHisasnGlyArgAspLeuGlnSerSerThrLeuSerValProGlu	1013
Db	2500	-----GAIAAAGAGAGAGGCAACAGAACCGCGCAAAACCGCGCGCCGAGGAGA	2554
Qy	1013	InValMetSerSerAsnHisCysSerProSerGlySerProHisArgValaspValileg	1033
Db	2555	AAACGAGAAAAAGC-----AAACCGCAAGAAAGGCCACAG	2593
Qy	1033	lvArgThr-ArgserTrpSerProSerAlapProProGlnArgasnValGluGlnGly	1052
Db	2594	GAGCACAGAAGAGGCAACAGGAGAAACCAAGCAAAAAAAGCAACACAGCAAGCAAAA	2653
Qy	1053	HisArgglyThrArgAlaThrGlyHisTyrrAsnThrIleSerArgMetAspArgHisArg	1072
Db	2654	CACAAA-----	2659
Qy	1073	ValMetAspAspHisTyrrSerSerAspArgAspArgAspCysGluAlaalaAspArgGln	1092
Db	2660	-----CCAAAGGACAGAGAGACAGAAAGGAGCCCAAGAA	2695
Qy	1093	Pro-----TyrrHisArgSerArgSerThrGluGln	1102
Db	2696	CCAACAGCAGCAAAAAGGAAGAGGGGAACAGAAACGACAGAGAGGGCAACAGGACAA	2755
Qy	1103	ArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgTrpProAspThrAsn	1122
Db	2756	AAACGAAGCAGGAAAAAAGCAAGAAACAGAGGGGGGCAAGAACCGCCAAAGCCCCAC	2815
Qy	1123	LeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerProAlaLeu	1142
Db	2816	GCNAAACAAAA---CCNAGGGGAAGACAGGACAGCGCGAGCAGACAGCGCGGAGCACA	2872
Qy	1143	SerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThr-----	1159
Db	2873	ACAAACAAACACAGAAAGCGGCGCAACGAAAAAGAAACAGAAAGAAACCCGAGGAGAA	2932
Qy	1160	-----ProGlyThrGlyArgGlyArgGlnLeuProGlnLeuPro	1173
Db	2933	CGNAAACGACACAAACGACGAGAGAGAGAAAGAAAGACAGAAAGAAACCCGAGAGCCCA	2992
Qy	1174	ProLysGlyThrLeuGluArgSerAlaMetAspileGluGluArgAsnArgGlnMetLys	1193

Db	2993	GC	AAAAA	AAAAA	AAAAA	AGAGGGGG	CAC	CAAG	CAGGGG	CAGCGGGG	CACAGAA	CAAA	3052	
Qy	1194	Leu	en	ly	s	Ty	Le	Val	Ala	Gly	Ser	Asp	Pro	Arg
Db	3053	GG	AA	CA	CG	ACCC	CA	CC	GGG	CAG	CA	AA	3112	
Qy	1208	Glu	-Gln	Asp	Ty	His	Ser	Le	Val	Tyr	Arg	Ser	Gly	T
Db	3113	GAC	ACA	GAC	AC	CA	CA	AG	AG	CC	CA	AA	3172	
Qy	1227	r	Val	Ser	Thr	Le	Ser	Ser	Asp	Ser	Asp	Val	Ser	Ala
Db	3173	AA	CG	AAAA	AA	CA	CG	CC	CA	AG	-----	CA	AA	3266
Qy	1247	r	Ser	Ala	Ser	Arg	Phe	Ser	Ser	Thr	Ser	Tyr	Met	Ser
Db	3227	CG	AA	GGG	CG	CA	AG	CAG	CG	CA	CA	CA	AA	3284
Qy	1267	Y	Asn	Arg	Gly	Ala	Ser	Val	Phe	Thr	Ser	Le	Ser	Gln
Db	3285	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Qy	1287	Y	Leu	Asn	Leu	Thr	Le	Ser	Ser	Thr	Ser	Le	Ser	Gly
Db	3290	AG	AG	CA	CC	CA	CG	AA	GG	CG	CG	CA	-----	CA
Qy	1307	p	Gly	Ser	Gln	Ser	Asp	Thr	Ala	Val	Gly	Ala	Leu	Gly
Db	3323	GA	AC	CAG	CA	GA	AA	AG	CA	CC	AA	AA	CG	AA
Qy	1327	r	Ser	-----	I	Leu	Ala	Le	Ser	Met	Val	Ala	Leu	Gly
Db	3382	AG	GA	AA	AG	AA	GG	CG	CG	CA	AA	-----	AG	AA
Qy	1345	r	Ala	Ser	Gln	Leu	Ser	Gln	Thr	Glu	Gly	Gly	Leu	Leu
Db	3421	AG	AA	CA	AA	CG	AA	AA						

RESULT 4
US-08-680-897-1/c
; Sequence 1, Application US/08680897
; Patent No. 6008051
; GENERAL INFORMATION:
; APPLICANT: DasSarma, Shiladitya
; APPLICANT: Halladay, John
; APPLICANT: Ng, Wai-lap
; TITLE OF INVENTION: RECOMBINANT VECTOR AND PROCESS FOR CELL
; TITLE OF INVENTION: FLOTATION
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/680,897
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/271,270
 ; FILING DATE:
 ; APPLICATION NUMBER: 944,581
 ; FILING DATE: 14-SEP-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Digiglio, Frank S.
 ; REGISTRATION NUMBER: 31,346
 ; REFERENCE/DOCKET NUMBER: 8680
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 742-4343
 ; TELEFAX: (516) 742-4366
 ; TELEX: 230 901 SANS UR
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9880 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-680-897-1

Alignment Scores:

Pred. No.: 9,6e-10 Length: 9880
 Score: 292.50 Matches: 356
 Percent Similarity: 33.48% Conservative: 178
 Best Local Similarity: 22.32% Mismatches: 582
 Query Match: 3.53% Indels: 480
 DB: 77 Gaps: 77

US-09-617-099B-1 (1-1590) x US-08-680-897-1 (1-9880)

QY 10 ArgProAlaPro---ThrProAlaLa-----SerGlnPro-ProProGlnProGluMe 26
 DB 4797 AGAATGCTCCAGGATCTCGGGCTGTTGGGGGAGACCTGAGTCCGGGTACAGTATA 4738
 QY 26 tProAspLeuSerHisLeuThrGluGluArgLysIleLeuAlaValMetAspAr 46
 DB 4737 CCGCATTTAAATGACCTTGACGCGAA-----GGTGTTACTTGAGGT 4696
 QY 46 gGlnLysLysGluGluGluLysGlnSerValLeuLysIleLysGluGluHisLysAl 66
 DB 4695 ACAGAAATTGAGCAAGCGCAA-----GTGTATCGCTCTCCGATCTCGAAGGAGC 4645
 QY 66 a-----GlnProThrGlnTrpPheProPheSerGlyIleThrGluLe 80
 DB 4644 GTTCAACCGAATAGACCATGTTGATGCTGCTGCTGTTTTCGCTGCTACTCAAGGC 4585
 QY 80 uValAsnAsnValLeuGlnProGlnLysGln-ProAsnGluLysGluProGlnThrL 100
 DB 4584 GGTATGACCGATTGCAAGCGGCTCAATCACAATCCCAAGGAAGAGACCAATGACTG 4525
 QY 100 yLeuHisGlnGlnPheGluMetTyrLysGlnValLysLysMetGlyGluGluSerG 120
 DB 4524 AGAACCTATACATACGATGATCATCGAACAGGAAGATCTCGAATAGATGTCGAAGGCG 4465
 QY 120 lnglnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysP 140
 DB 4464 TTGCGCGAGCGGAACAG-----GTCTATACGCTCGATTACAGACACTCT 4420
 QY 140 heAlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysG 160
 DB 4419 CCGCT-----GTGCTCTCTGACATTTGATACGACCGACCCCGAGCGACCG 4375
 QY 160 lyGlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysG 180
 DB 4374 ATGAGGACGTGGGAAGCTCATAACACGCTCTCAGGAGGATTGGAAGCACGAGAGAAC 4315

QY 180 lnglnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuG 200
 DB 4314 GCACAGTCTCCGATGAGCTTCGGGATGGCTTC-----AAAAGTCCGGCGACGTAA 4261
 QY 200 lnglnProAspGlnLysValProArgGlyLeuArgAsnGlu-----GluAlaP 216
 DB 4260 AGGGTGATTTCGGCGGGCGGACGTGCAATTTCGACGATGACCTGAATGACATCGAGGAA 4201
 QY 216 roGlnGluLysLysAlaLysLeuHisGluGlnProGlnPheGlnGlnGlnAlaProGlyAspL 236
 DB 4200 CCGTCGAACCTTGGCGTGAAGATACTC-----GGTCCTGCGGACG 4162
 QY 236 euSerValProAlaValGlu---LysGlyArgAlaHisGlyLeuThrArgGlnAspThrI 255
 DB 4161 ATACAGTCCCTCGAGAAAGAAATCCAGGAAACGTCACGATCACTACAGACA----- 4110
 QY 255 eLysAsnGlySerGlyValLysHisGlnIleAlaSerAspMetProSerAspArgLysAr 275
 DB 4109 -----TCTGAGTATCAACGAGACCGAGAACG 4084
 QY 275 gser-----ProSerValSerArgAsp----- 282
 DB 4083 ATCTCTTCACAGACCGCTGATCATATAAGTGTACCTCTGCTGCTGCTGAGAAACGCG 4024
 QY 283 -----GlnAsnArgArgTyrGluGlnSerGluGluArgGluAsp----- 295
 DB 4023 ATGCTTTCGACTCCGCCCATCGACGATGTCGAAGCGGAATACGACCACTGACGATCAGT 3964
 QY 296 -----TyrSerGlnTyrValProSerAspGlyThrMetPr 307
 DB 3963 ACACGGGGCGTGGCGCGGTACAAATTCGTGGACATTCACATCGCGCGGACACAGCAAC 3904
 QY 307 oArgSerProSerAspTyrAla---AspArgArgSerGlnArgGluProGlnPhe----- 324
 DB 3903 AAGGAGCGCGATATGATCATGATGTCATCATAGACGATCTCTCTGTCGAGCCGTTTTCTCT 3844
 QY 325 -----TyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgGlyHisAr 342
 DB 3843 TGCTCGATATCTCCAGAC-----GATGGCTTCACGAGATGTAGC 3802
 QY 342 gHisSerLysGluTyrIleValAspAsp-GluAspVal---GluSerArgAspGluTyrG 361
 DB 3801 ACACCAA-----ATCGATTCGAGACGACATCAAGGAGAACCACTCTCTGTAGC 3754
 QY 361 lu---ArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuA 380
 DB 3753 AAGTTGGTGAACGATCCGACGAGGAGTACCAAGCAAGATGCGTACCGCATGGAGGTCA 3703
 QY 380 laArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValS 400
 DB 3702 AATCGCAACTGGAGATGGCGGAGCGGTCCAAAGAACAGATGCGTACCGCATGGAGGTCA 3643
 QY 400 exArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluA 420
 DB 3642 AGAATAATGGTACCCGACGAAA-----ACGACGACGGGTCCGACGACCATCTTCCCA 3589
 QY 420 spSerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluA 440
 DB 3588 GCTCAGCGGCTGCTCGACCAATTCACACGCTGCTCGAGGTGCTCGCGCATCGAACA 3529
 QY 440 rArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnG 460
 DB 3528 GGAGGGCGGTACCGACACGATGAGGCGGTATCGACCGGGGAAACCGCCGATTTGATTA 3469
 QY 460 lyGlnSerSerTyrProGln-----ArgThrSerAsnH 471
 DB 3468 CGACTACGACGTTTCCATTCGGCTGGACGACGCGGCTGCTCTACGACGAGGAAC 3409
 QY 471 iserPro-----ProThrProArgSerProIleProLeuAspArgProAspMetA 489
 DB 3408 ATCGTCCCAATCGTCCCGCTCCGACAGCGATCCCGTTCAGCAAGAGCA-----C 3358
 QY 489 rArgAlaAspSerLeuArgLysGlnHisLeuAspProSerSerAlaValArgLysTr 509

Db 3357 GGAGG-:::|||||ATTGATTCA 3343
QY 509 hrLyAsrGluLyMetGluThrMetLeuArgAsnAspSerLeuSer-:::||||| 524
Db 3342 CGTCGAGACTCGTGGAGACAGCGCGGACCACTCGTCTGCTGCGAGTTGCGCGG 3283
QY 525 -----SerAspGlnSerGluSerValArgProProProProArgProHisL 540
Db 3282 CGTGACGGACGACGACGTCGACGTCGCACTCGATACCG-:::||||| 3245
QY 540 ysSerLySylsGlyGlyMetArgGlnValSerLeuSerSerSerGluGluLeuA 560
Db 3244 -----ACGAGGAGCACTCAAGCTACG 3223
QY 560 laSerThrProGluThrSerCysAspValGluLeuLeuGluSerGluSerValSerG 580
Db 3222 CGTCGACGACGACGTCGTCGAC-:::|||||GATCGTACTCGACCGGCC 3181
QY 580 lu-lysgly-----AspSerGlnlys---GlyLyAsrGlyThrSerGluGln 594
Db 3180 GGAGCTGGCATCACCGATATGACTCTCAGAAATCAGTTCTCGAAATTCGGCTAGCACG 3121
QY 595 GlyVal-----LeuSerAspSerAsnThrArgSerGluArg 606
Db 3120 AACATCGGACACCGATGGAGTGAATCCATGACGCGCAACAAACAGCAAAACACAG 3061
QY 607 GlnLySylsArgMetTyrTyrGlyGlyHisSerLeuGluLeuAsp-:::||||| 621
Db 3060 CAGAAACGACAGGACGAGTAAAGCGGACGATCAACCGGATAAAGCCCGCGCGCAAT 3001
QY 622 -----LeuGluThrTrpSerGlu 626
Db 3000 CTGCTCGGACGCGGAAACTCGCCGCGTGGACCGCAACCGGAACAGTCGAA 2941
QY 627 ProGlnIleLyAspSerGlyValAspThrCysSerSerThrThrLeuAsnGluHis 646
Db 2940 CCCCGCGTGACGATTCGACGCGGACGACTACTGACGCGGAGACGGTCAAGAAATCCAGCG 2881
QY 647 SerHisSerAspLyHisProValThrTrpGluProSerLyAspGlyAspArgLeulle 666
Db 2880 GCTCAGTCG-----ACGATCCCGCTCAGAAAGTCGAACGCGGAG----- 2842
QY 667 GlyArgIleLeuLeuAsnLyAsrGluLeuLyAspGlySerValProArgAspSer----- 684
Db 2841 -----AACCGGTGACCACTCCCACTCGACCGTCCGAGACTCGAAGTACTCGATGTC 2785
QY 685 -----GlyAlaMetLeuGlyLeuLyValValGlyGlyLyMet 697
Db 2784 ACGGCTCGAGCGGATGTACGCGCAACGCTTCACCAAGAAACAGCGCGTGGAAACTCG 2725
QY 698 ThrGluSerGlyArgLeuCysAlaPheIleThrLySylsValLySylsGlySerLeuAlaAsp 717
Db 2724 GAGGCATCCGCC----- 2713
QY 718 ThrValGlyHisLeuArgProGlyAspGluValLeuGluThrPAsnGlyArgLeuLeuGln 737
Db 2712 -----CGTCGATGAG-----ACGCGTCC 2692
QY 738 GlyAlaThrPheGluGluValTyrAsnIleIleLeuGluSerLySylsProGluProGlnVal 757
Db 2691 GGGAGTCCCACTGAGGACGAGGTGAAC-----GATGAGTGACCCCAACC 2647
QY 758 GluLeuValValSerArgPro-----IleGlyAspIleProArgIleProAspSer 774
Db 2646 GACGCGCTCGCAGGCGGACCTCGCCGAGATGCTGGAGATGCTCTCTAGACA-AAGGAGTCG 2588
QY 775 ThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGlnLySylsMetAspArgPro 794
Db 2587 TCTCAACGGGATATCGCAGTCAGCGTCGGGACACGGAAC----- 2546
QY 795 SerIleSerValThrSerProMetSerProGlyMetLeuArgAspValProGlnPheLeu 814
Db 795 -----ACGAGACGAGACCTCGAACAGGT-----GAAGCGATGGTGGTCACGAC 1724

2545 -----TCCTCGGTATCGAGTTACGGCCCGGATTGCTTCTGTCGACAGCGCGCAAT 2492
QY 815 SerGlyGlnLeuSerIle-----LysLeuThrPhe-----AspLysValGly 828
Db 2491 ACGGCTCGAGTTCCCAACCGGACGATATGAGCGCGTCGAGTCGCGCGCAATATCT 2432
QY 829 HisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuProSerArgGluAspGlyArg 848
Db 2431 CACGGACCACTCGGACC-----CGG 2411
QY 849 ProArgAsnProTyrValLySylsIleTyrPheLeuProAspArgSerAspLysAsnLyAsr 868
Db 2410 CGTCGAAA-----CCCAATCGAGACCGAGTCGACGAAAC 2375
QY 869 ArgThrLySylsValLySylsThrLeuGluProLySylsTrpAsnGlnThrPheIleTyrSer 888
Db 2374 CACTGTGGACGATTGACGCGCGCATCCA----- 2342
QY 889 ProValHisArgArgGluPheArgGluArgMetLeuGluIleThrLeuTrpAspGlnAla 908
Db 2341 -----CGTCGCGGAGGACCAATGAGATGAGTACTGACTCGACGACGCG 2291
QY 909 ArgValArgGluGluSerGluPheLeuGlyGluIleLeuIleGluLeuGluThrAla 928
Db 2290 GACGATTTCAGAGCGGCTCACGCGCTGCTGTAATCTGCTCGAATTCCTGCTCGAG 2231
QY 929 LeuLeuAspAspGluProHisTrpTyrLySylsLeuGlnThrHisAspValSerSerLeuPro 948
Db 2230 GCCCTCGAACAGAA-----GCGGTGCTCGGATGGAATCCGCGTCTGCTC--- 2186
QY 949 LeuProArgProSerProTyrLeuProArgArgGlnLeuHisGlyGluSerProThrArg 968
Db 2185 -----TCAGAGGACGAAATCGAA 2168
QY 969 ArgLeuGlnArgSerLyAsrGlySerAspSerGluValSerAspTyrAspCysGlyAsp 988
Db 2167 CGATTGCGGCGCAATACAGGCTCGAAGACGAGTCTCGAGCGACTCGAAGACGACGAG 2108
QY 989 GlyValGlyVal-----ValSerAspTyrArgHisAsnGlyArgAspLeuGlnSerThr 1007
Db 2107 GATATCAACGCGCGAGTCTCCGAGTTAGG-----GAGGACCTCGATCATCGTGATC 2057
QY 1008 LeuSerValProGluGlnValMetSerSerAsnHisCysSerProSer-----GlySer 1025
Db 2056 CCGGACCGCATCGAACAGCTA-----TCCGACGACGACACCGCTCAGCTCAGGCTCG 2003
QY 1026 Pro-----HisArgValAspVal-----IleGlyArgThrArgSerTrpSerPro 1040
Db 2002 CCGAATCCCGAGCTGATGACGTATGACTGACCCCGGCCAGCCCGGAGAGAGAGC--- 1946
QY 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
Db 1945 -----AGAACACAGCGAACG-----AGGAACGCGACGCTCAGCAACGCGC 1907
QY 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
Db 1906 GCTATCTTACTCGGTGTCGATACCACTCGTCTGCG----- 1871
QY 1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
Db 1870 -----AATCGCGACCTGTCTCCAGCACCGGGTCTG--- 1841
QY 1101 GluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProAsp 1120
Db 1840 -----ACGACAACTGTCTTACCTGCTG---AGGCGC--- 1811
QY 1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140
Db 1810 -----ATGCGTGGCGCGCTGCTCATGACTGTGAGACGGTCT--- 1772
QY 1141 AlaLeuSerArg-SerHisProArgThrGlySerValGlnThrSerProSerSerThrPr 1160
Db 1771 -----ACGAGACGAGACCTCGAACAGGT-----GAAGCGATGGTGGTCACGAC 1724

QY 1160 oGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArg 1180
Db 1723 AGCAGGTGTCACCGCGC----- 1705
QY 1180 gSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLysLeuAsn-LysTyrLysGlnV 1200
Db 1704 -----GAGCGAGCGGTCGTGACCGCGTCGCGATCGGATTCGACGCG 1661
QY 1200 al---AlaGlySerAspProArgLeuGlu-----GlnAspTyrHisSerLysTyrA 1216
Db 1660 TCCTCGAGCGCGGTGATGCGAGTATCGAAGCGGTGTAGAACACCACTACGAGGCGCTCC 1601
QY 1216 rgSerGly-----TpaSp- 1220
Db 1600 GCGAGCAATACGTCGTCGCGGAGTGTGGAGTATCGAATCAATCTGTGTGGATT 1541
QY 1221 ProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerValSerVal 1240
Db 1540 CGCGACCGTTCGAGGAGCATCGCAGACCGGACCGCGCTCCGAGAACTACGACG 1481
QY 1241 SerAlaValSerArgThrSerSerAlaSerArgPheSer-SerThrSerTyrMetSerVa 1260
Db 1480 GCGAGCAACATCGCGCGCAGGAGAAAGTCTCTCTCGAGAAACA-----GT 1433
QY 1260 lGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAs 1280
Db 1432 CGCATCAGCATCTCAAGAGCTGAACGAGAGCGCGGACGAGTACGATCACTGA 1373
QY 1280 nArg-GlnMetGlyValSerGlyLysAsnLeuThrLys-----SerThrSerLys 1297
Db 1372 AGAGGCCATTACCCCGTCGTGACGACCTGACCGACGAGCAGCATACGCGCTAC 1313
QY 1297 exGlyAspMetCysSerLeuGluLysAsnAsp-----Glys 1309
Db 1312 AGGACGACACATCGTCATCGAAGAAACAGACATCGTCGGTTCGCGCTTCGCGGAG 1253
QY 1309 exGlnSerAspThrAlaValGly---AlaLeuGlyThr-SerGlyLysLysArg----- 1325
Db 1252 AGGACGACGAGACCGCTCTCGGTGATCGATTGGATACGATCTCGAACACGAGGGGTAG 1193
QY 1326 ArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSer 1345
Db 1192 AGATCAGAT-----TCAGCGGCGCGTGGCCACCGTACA 1160
QY 1346 AlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArg 1365
Db 1159 CGTTCGCGCCAGATATTGGTAAATATGAGCCAAACAAAGACGACACACGCGATCT 1100
QY 1366 SerThrGluThrGlyLeuAlaValGluMetArgAsnTyrMetThrArgGlnAlaSerArg 1385
Db 1099 -----TCAGTTCGTCGACGCTTACTGCGGCGGAGCGCT 1064
QY 1386 GluSerThr-----AspGlySerMetAsnSerTyrSerSerGluGlyAsnLeu 1401
Db 1063 GATTCAAGCGGCGATGATCGTACGCTCGCCGACATTCCTTCATCGGATCAGCCT--- 1007
QY 1402 IlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLeu 1421
Db 1006 -----CCGGC-----AGCGATTG---TGGCATGACCCATCAGCGAGTA 968
QY 1422 GlyProAlaGlnLeuValGly 1428
Db 967 CGGCTGTCGAGGAGTGGA 947

RESULT 5

US-08-961-527-363

; Sequence 363, Application US/08961527

; Patent No. 6420135

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 363:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4483 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-363

Alignment Scores:

Pred. No.: 4,62e-10 Length: 4483
Score: 290.00 Matches: 295
Percent Similarity: 33.68% Conservative: 221
Best Local Similarity: 19.26% Mismatches: 632
Query Match: 3.50% Indels: 388
DB: 4 Gaps: 61

US-09-617-099B-1 (1-1590) x US-08-961-527-363 (1-4483)

QY 7 ProArgGlyArgProAlaPro-----ThrProAlaAlaSerGlnProPro 22
Db 324 COTCAGCATCGACAGCGCTCAGCTTCAGCAGTACCGAGTCTTCAGCCTCAGCGTCA 383
QY 23 GlnProGluMetProAspLeuSerHisLeuThrGluGluGlu-----Arg 37
Db 384 CAGTGGCTCGCGCTCAACCGATGATCTGATCGGCATCAACCGTGGTGGTCTCAG 443
QY 38 LysIleIleLeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerVal 57
Db 444 CAAGTACTAGTGCATCAGCTTCAGCATCAACGAGTGCATCGGCTTCAGCATCAACGAGT 503
QY 58 LeuLysIleLysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIle 77
Db 504 COTCGCTTCAGCGTCAACCGATGCGTCAAG---CTTCAGCAAGTACCGATG---CTT 554
QY 78 ThrGluLeuValAsnAsnValLeuGlnProGlnLysGln---ProAsnGluLysGlu 96
Db 555 CAGTCTCAGCATCAACAGTCTTCAGCTCAGCATCGACGATCGCGCTTCGGCTTCAGCA 614
QY 97 ProGlnThrLysLeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGly 116
Db 615 GCACATCAGCATCTGATCAGCGTCAACGCGCTCAGCTTCAGCAAGTACCGATGCGT 674
QY 117 GluGluSerGlnGlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHis 136
Db 675 CAGCTCAGCGTCGACAAAGTGGTCAAG---CCTCAGCAAGTACTAGTGCAT 722
QY 137 LysThrLys-PheAlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheC 156

```

723 CAGCTTCAGCATCAACGAGTGCATCGCTTCGGCGTCAACGAGTGCATCAGAGTCAGCAA 782
156 sAlaArgCysGlyGlyArgValSerLeuArgSerAsnLysValMet---TrpValCysAs 175
783 GTACCAAGTGC-----GTCAGCTTCGGCATCAACAGTGCCTCGGCTTCAGCAAGCA 833
175 nLeuCyArgLysGlnGlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsnSerG1 195
834 CCAGTGGCT-CGGCTTCAGCAAGTACTAGCGCT-----CAGCCTCA 874
195 ySerAsnThrLeuGlnGlnProAspGlnLysValProArgGlyLeuArgAsnGluGlu-- 214
875 GCCTCAACCAAGTGGCTCAGCCTCAGCAAGTATCTCAGCGCTGTAATCGGCATCAACGAGT 934
215 -----AlaProGlnGlnLysLysAlaLysLeuHisGlu----- 225
935 GCGTCCGCTTCAGCAAGTACTAGCGCTCAGCCTCAGCGTCAACAGTGCATCGGCTTCA 994
225 ----- 225
995 GCGTCAACGAGTGGCTCTGAATCGGCATCAACGAGTGGCTTCAGCAAGTACTAGC 1054
226 ---GlnProGlnPheGlnGlyAlaPro-----G1 234
1055 GCCTCAGCCTCAGCGTCAACAGTGCATCGGCTTCAGCATCAACGAGTGGCTTCGCTTCA 1114
234 yAspLeuSerValProAlaValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspTh 254
1115 GCAAGTACTAGCGCTCAGCCTCAGCGTCAACAGTGCATCGGCTTCAGCGTCAACGAGT 1174
254 r---IleLeuAsnGlySerGlyValLys-----HisGlnIleAlaSerAs 268
1175 GCGTCTGAGTCAGCATCAACGAGTGGCTCAGCCTCAGCAAGCATCAGCTTCT-GAATC 1233
268 pMetProSerAspArgLysArgSerProSerValSerArgAspGlnAsnArgTyrG1 288
1234 TGCATCAACAGTGGCTCAGCCTCAGCATCGCAAGCGCCTCAGCTTCAGCAAGTACTACAG 1293
288 uGlnSerGluGluArgGluAspTyrSer-----GlnTyrValProSerAspGlyThrMe 306
1294 TGGTTCAGCTCAGCGTGCAGCAAGTGCCTCGCTTCAGCAAGTACTACAGTGC-----CGTCA 1347
306 tProArgSer---ProSerAspTyrAlaAspArgArgSerGlnArgGluProGlnPheTy 325
1348 GCCTCAGCAAGTACCAGTGCTCAGCCTCAGCGTTCGACAGTGGCTCGGCTTCA----- 1402
325 rGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGlyHisArgHisSerLy 345
1403 -----CCAGTGCATCTGAATCGGCATCAACAGTGGCTCAG----- 1438
345 sGluTyrIleValAspAspGluAspValGluSerArgAspGluTyrGluArgGlnArgAr 365
1439 -----CCTCAGCAAGTACTAGCGCTCAGCCTCAGCATCAACAGTGGCTGC 1485
365 gGluGluGluTyr-----GlnAlaArgTyrAr 374
1486 GCTTCAGCAAGTACTAGTGCATCAGCTTCAGCAAGTACTAGCGCTCAGCTCAGCGTGC 1545
374 gSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMetAr 394
1546 ACAAGCGCTTCAGCTTCAGCAAGTACCAGTGGCT---CAGCCTCAGCGTTCGACAGTGGC 1602
394 g-----IleHisAlaGluValSerArgAlaArgHI 404
1603 TCGGCTTCAGCAAGTACTCAGCGTCTGAATCAGCATCAACAGTGGCTCGGCTTCAGCA 1662
404 sGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAspSerArgIleSe 424
1663 TCACCAAGTGCAT-----CAGCTTCAGCATCAACAGTGGCT 1698
424 rLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArgArgAlaAlaMe 444
1699 TCAGCTT-----CAGCAAGTACTACAGTGGCTCGG----- 1726
444 tGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGlyGlnSerSerTy 464
1727 -----CTTCAGCATCAACGAGTGCCTCAGTCTCAG----- 1756
464 xProGlnArgThrSerAsnHisSerProProThrProArgSerProIleProLeuAs 484
1757 -----CGTCAACCAAGTGGCTCTGAA----- 1776
484 pArg-----ProAspMetArgArgAlaAspSerLeuArgLysGlnHisLe 500
1777 TCGGCATCAACAAGTGCCTCGGCTTCAGCAAGCAGCAGTCTTCGGCTTCAGCGT----- 1831
500 uAspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAs 520
1832 -----CAACGAGTGGCTCTGAGTCAGATCAACGAGTGGCTCAGCCTCAGCAAGCA 1881
520 n-----AspSerLeuSerSerAspGlnSerG1 529
1882 CATCAGCTTCTGAATCTGCATCAACGAGTGGCTCCTTCGCGCATCAACAGCGCTTCGGC 1941
529 uSerValArgProProProArgProHisLysSerLysLysGlyLysMetArgG1 549
1942 CTCAGCA-----AGTACAAGTGGCTTCAGCCTCAGCATCAACAGTGCATC 1986
549 nValSerLeuSerSerSerGluGluLeuAlaSerThrProGluTyrThrSerCysAs 569
1987 AGCTTCAGCCTCAACAGTGGCTTCAGCCTCAGCGTCAACGAGTGGCTTCGCGCTTCAGCAAG 2046
569 pAspValGluLeuGluSerGluSerValSerGluLysGlyAspSerGlnLysGlyLysAr 589
2047 TACCAGTGGCTCAGCTTCAGCAAGCAAGTGGCTCAGCTTCAGCATCAACAGTGCCTTC 2106
589 gLysThrSerGluGlnGlyValLeuSerAspSerAsnThrArgSerGluArgGlnLysLy 609
2107 GGCTTCGGCATCAACAAGTGGCTTCAGCATCAGCATCAACGAGTGGCTCAGCAAGT 2166
609 sArgMetTyrTyrGlyGlyHisSerLeuGluLeuAspLeuGluTyrSerGluProGlnI1 629
2167 ACTAGTGCATCAGCATCAGCT-----CAACAGTGC 2199
629 eLysAspSerGlyValAspThrCysSerSerThrThrLeuAsn-----G1 644
2200 TCAGCCTCAGCAAGTA-----TCTCAGCTCTGAATCGGCATCAACAGTGC 2247
644 uGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLysAspGlyAspAr 664
2248 TCAGCATCAGCATCAACGAGTGCATCGGCTTCAGCGTCAACAGTGCATCAGTCTCAGCA 2307
664 gLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAspSe 684
2308 AGCACCAGTGGCT-----CGGCTTCAGCATCAACAGTGGCTCAGCGCT-- 2350
684 rGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSerGlyArgLeuCy 704
2351 -----CAGCAAGTACTCAGCTCTGAATCGGCATCAACGA----- 2386
704 sAlaPheIleThrLysValLysGlySerLeuAlaAspThrValGlyHisLeuArgPr 724
2387 -----GTGCGTCA 2394
724 oGlyAspGluValLeuGluTyrAsnGlnGlyArgLeuLeuGlnGlyAlaThrPheGluGluVa 744
2395 GCCTCAGCAAGTA-----CTAGTGCATCAGCATCAGCATCAACAGT 2436
744 lTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuValValSerArgPr 764
2437 GCATCG-----CTTCAGCAAGTACCAGCGCTTCAGTCTCAG-----CAAGCACC 2481
764 oIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGluSerSerSerSe 784
2482 AGTGGCTCAGCTCAGCAAGTACCA----- 2506
```

QY 784 rSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSerProMetSerPr 804
DB : : : : :
DB -GCGCTCAGCTCAGCAAGCAGCAGTGGCTCAGCTTACAGCAAGTACCAAGTGGTTCAGCC 2565
QY 804 o-----GlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSe 819
DB : : : : :
DB TCAGCGTCGCAAGTGGCTTCAGCAAGTGGCTTCAGCTTGAATCAGCATCAAG 2625
QY 819 rIleLysLeuThrPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAla-- 838
DB : : : : :
DB 2626 AGTG-----CATCAGCTTCAGCATCAACAAGTGGCTTCAGCTTCAGCAAGTACC 2673
QY 839 -----LysAspLeuProSerArgGluAspGlyArgProArgAs 851
DB : : : : :
DB 2674 AGTGGTGGCTTCAGCATCAAGTGGCTTCAGCTTCAGCTTCAGCAAGTGGCTTCGAA 2733
QY 851 nProTyrValLysIleTyrPheLeuProAspArgSerAspLysAsnLysAlaArgThrIly 871
DB : : : : :
DB 2734 TCAGCATCAACAAGTG-----CCTCGGCTTCAGCAAGTGGCTTCAGCTTCAGCAAGTGG 2784
QY 871 sThrValLysLysThrLeuGluProLysTrpAsnGlnThrPheIleTyrSerProValHi 891
DB : : : : :
DB 2785 G-----CAAGTACTAGTGCATGGCTTCAGCA 2811
QY 891 sArgArgGluPheArgGluArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValAr 911
DB : : : : :
DB 2812 TCAGCAAGTGGCTTCGAATCGGCATCAACAGTGGCTTCAGCTTCAGCAAGTGG 2871
QY 911 gGluGluGluSerGluPheLeuGlyGluIleLeuIleGluLeuGluThrAlaLeuLeuAs 931
DB : : : : :
DB 2872 TCAGCTCAG----- 2881
QY 931 pAspGluProHisTrpTyrLysLeuGlnThrHisAspValSerSerLeuProLeuPro-- 950
DB : : : : :
DB 2882 -----CAAGCATCATCAGCTTCTGAATCTGCATCAACCACT 2916
QY 951 -ArgProSerProTyrLeuProArgArgGlnLeuHisGlyGluSerProThrArg--Ar 969
DB : : : : :
DB 2917 GCGTGGCTTCAGCGTCAACCAAGTGGCTTCAGCTTCAGCTTCAGCAAGTGGCTTCG 2976
QY 969 gLeuGlnArgSerLysArgIleSerAspSerGluValSerAspTyrAspCysGluAspGl 989
DB : : : : :
DB 2977 GCATCAACAGTGGCTTCAGCGCTTCAGCAAGTGGCTTCAGCTTCAGCAAGTGGCTTCG 3034
QY 989 yValGlyValValSerAspTyrArgHisAsnGlyArgAspLeuGlnSerSerThrLeuSe 1009
DB : : : : :
DB 3035 -----CAACAGTGGCTTCGCTTCAGCAAGCAGCAAGTGG 3069
QY 1009 rValProGluGlnValMetSerSer--AsnHisCysSerProSerGlySerProHisAr 1028
DB : : : : :
DB 3070 TCAGCTCAGCAAGTATCTCAGCGTCTGAATCGGCATCAACAGTGGCTTCAGCA 3129
QY 1028 gValAspValIleGlyArgThrArgSerTrpSerProSerAlaProProGlnArgAs 1048
DB : : : : :
DB 3130 AGCGCAAGTACTCAGCGT-----CAGCTTCGCGCTCAACCAAGTGGCTTCG 3174
QY 1048 nValGluGlnGly-----HisAr 1054
DB : : : : :
DB 3175 GCTTCAGCAAGCAGCAAGTGGCTTCAGCTTCAGCAAGTATCTCAGCTTCGAATCGGCATCA 3234
QY 1054 gGlyThrArgAlaThrGlyHisTyrAsnThrIleSerArgMetAspArgHisArgValMe 1074
DB : : : : :
DB 3235 ACAGTGGCTTCAGTTCAGTTCAGCATCAACAGTGGCTTCAGCTTCAGCAAGCAGCATCAGCTTCT 3294
QY 1074 tAspAspHisTyrSerSerAspArgAspCysGluAlaAlaAspArgGlnProTy 1094
DB : : : : :
DB 3295 GAATCGGCATCAACCAAGTG-----CGTCAGCTCA 3324
QY 1094 rHisArgSerArgSerThrGluGlnArgProLeuLeuGluArgThrThrArgSerAr 1114
DB : : : : :
DB 3325 GCATCGA-----CAAGCGCTCAGCTTCAGCA-AGTACCAAGTGGCTTCAGC 3368

QY 1114 gSerSerGluArgProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyAr 1134
DB : : : : :
DB 3369 CTCAGCGTCGCAAGTGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 3428
QY 1134 gSerAlaProProSerProAlaLeuSerArgSer-HisProArgThrThrGlySerValGlnT 1154
DB : : : : :
DB 3429 CTCAGCTTCAGCAAGTACTAGTGCATCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 3488
QY 1154 hrSerProSerSerThrProGlyThrGlyArgArgGly--ArgGlnLeuProGlnLeuP 1173
DB : : : : :
DB 3489 ATCAACCA-----GTGCTCGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 3542
QY 1173 roProLysGlyThrLeuGluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetL 1193
DB : : : : :
DB 3543 CAGTG-----CITCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 3572
QY 1193 ysLeuAsnLysTyrLysGlnValAlaGlySerAspProArgLeuGluGlnAspTyrHis 1213
DB : : : : :
DB 3573 TTCAGCTTCAGCATCGACAAGTG-----CCTCGGCTTCAGCAAGCAGCATCAGC 3620
QY 1213 erLysTyrArgSerGlyTrpAspProHisArgGlyAlaAspThrValSerThrLysSers 1233
DB : : : : :
DB 3621 ATCTGAATCAGCTCGACAGCGCTCAGCT-TCAGCAAGTACCAAGTGGCTTCAGCTTCAG 3679
QY 1233 erAspSerAspValSerValSerAlaValSerArgThrSerSerAlaSerArgPheS 1253
DB : : : : :
DB 3680 CGTCAGCAAGTGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 3739
QY 1253 erSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleSerV 1273
DB : : : : :
DB 3740 CATCGCTTCGCGTTCACCAAGTGCATCAGATCAGCAAGTACCAAGTGGCTTCAGCTTCOG 3799
QY 1273 alPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLys- 1292
DB : : : : :
DB 3800 CATCAACAAGTGGCTTCAGCTTCAGCAAGCAGTGGCTTCAGCTTCAGCTTCAGCTTCAG 3859
QY 1293 -----SerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspG 1308
DB : : : : :
DB 3860 CCTCAGCTTCAGCTTCACCAAGTGGCTCA-----GCCTCAGCAAGTATCTCAGCGT 3910
QY 1308 lysSerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgSers 1328
DB : : : : :
DB 3911 CTGAATCGGCATCAACAGTGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 3970
QY 1328 erIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerg 1348
DB : : : : :
DB 3971 CAACAAGTGCATCGGCTTCAGCTTCAGCAAGCAGTGGCTTCGAATCGGCATCAACAGTGGCT 4030
QY 1348 inLeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerThrG 1368
DB : : : : :
DB 4031 CGCTTCAGCAAGTACTAGCGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 4090
QY 1368 luThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerT 1388
DB : : : : :
DB 4091 CAACAGTGGCTTCGCTTCAGCTTCAGCAAGTACTAGCGCTTCAGCTTCAGCTTCAGCTTCAG 1398
QY 1388 hrAspGlySerMetAsnSerTyrSerSerGlu 1398
DB : : : : :
DB 4148 CATCGGTTTCAGCGTCAACAGTGGCTTCAG 4179

RESULT 6

US-09-144-085-3/c
; Sequence 3, Application US/09144085
; Patent No. 6280999
; GENERAL INFORMATION:
; APPLICANT: Gustafsson, Claes
; APPLICANT: Betlach, Mary C.
; APPLICANT: Ashley, Gary
; APPLICANT: Julien, Bryan
; APPLICANT: Ziermann, Rainer
; TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 30062-20020.20

```

; CURRENT APPLICATION NUMBER: US/09/144,085
; CURRENT FILING DATE: 1998-08-31
; EARLIER APPLICATION NUMBER: 09/010,809
; EARLIER FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 3
; LENGTH: 33529
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-144-085-3

```

Alignment Scores:	1.01e-08	Length:	33529
Pred. No.:	Score:	Matches:	348
	288.50	Conservative:	165
Percent Similarity:	31.55%	Mismatches:	581
Best Local Similarity:	21.40%	Indels:	534
Query Match:	3.48%	Gaps:	72
DB:	3		

US-09-617-099B-1 (1-1590) x US-09-144-085-3 (1-33529)

QY	2	SerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnProPro	21
Db	15682	TCAGGCCGACGGGGCAAGC---CAGCCCGCTCCAGTGTCCACGCCGTCTCCAGCA	15626
QY	22	Pro-----GlnProGluMetProAspIeuSerHisLeuThr-GluGluGluAr	37
Db	15625	CCAGCGCTGTGGGATCCATCGCTGTGCGCTCTCGGGCGGATATCCGAGAAACCG	15566
QY	37	g-----Ly	38
Db	15565	CATCGAACAGGTGATGTCCCGCAGGAATCCACCTTCGCGCACGTAGTCTTGTCCACCG	15506
QY	38	silelleLeuAlaValMetAspArgGlnLysLysGluGluGluLysGlu-----GlnSe	56
Db	15505	CCTCGGATCGGGGTGTAATCGAAGGTGCATCCAGCGGCGCAGGAAAGCTCGATCG	15446
QY	56	rValLeuLyslleLysGluGlu-----	63
Db	15445	CATCTCTCCCTCTCCAGGAGACGCCAGTAGTCTTCCGGCGTAGCAGCGCCCCCGCA	15386
QY	64	-----HisLys-----Al	66
Db	15385	GCCGCACGCCCATCATAGATGCGGTGGCCCTGAACGGAGATCCGAGTGGCTGCCT	15326
QY	66	agLlnProThrGlnTrpPheProPheSerGly-----IleThrGlu-----	79
Db	15325	GCAGCCTCTGCAACTCTTTTCAATGCACGGGTCAACGCCTCAAGAAAGCGTAGGAGAGCGG	15266
QY	80	-----Leu-ValAsnValLeuGlnProGlnGlnLysGlnP	92
Db	15265	CGCTCTATTATGAGAGATCTCTGAGTTCGAATCATTTCCAGGGCTTGCA	15213
QY	92	roAsnGluLysGluProGlnThrLysLeuHisGlnGlnPheGluMet-----	107
Db	15212	-----GGCCAGCTGCTCGTCGCGAGATCTTCCAAATCGAAATCGGAAC	15167
QY	108	-----TyrLysGluGlnValLysLysMetGlyGluGlnSerG	120
Db	15166	GGAACCTCGACGGATGTTTCGACGGCCCCCGCAGCTTCAAGAGCTTGGGAGAGCCCT	15107
QY	120	lInGlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysP	140
Db	15106	GAAACGACGACGAGCTACGGGAGGGCTGAGCAT-----GAACGCCGCAATT	15059
QY	140	heAlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysG	160
Db	15058	TCTCGTCGACATCGAGTCCGCTGTACG-----	15030
QY	160	lyGlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysG	180
Db	15029	-----CAGCAGATTCCGGCGCTCACCGTTTCCGAATCT-----TGAC	14993

QY 483 uAspArgProAspMet-ArgArgAlaAspSerLeuArgLysGlnHisHisLeuAspProS 503
Db 14054 AGCTCGCGGCCCCCGCGGCTCCAGCCGCGCGCTGACAGCGCTGACAGAGTGCAGCACC 13995
QY 503 erSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsnAspSerL 523
Db 13994 CCGTGGCGCGCAGGTCGCGCGGAGCGGCTGACCCAGCTC----- 13951
QY 523 euSerSerAspGlnSerGluSerValA-GrProProProArgProHisLysSerLysL 543
Db 13950 -----CCCTGTCCCGCGCTCACCAGCAC----- 13927
QY 543 ysGlyGlyLysMetArgGlnValSerLeuSerSerGluGluGluLeuAlaSerThrP 563
Db 13926 -----CGTCCGCGAGGTCAGCTCGCGGCTCGGTGAGCTCTTC----- 13885
QY 563 roGluTyrThrSerCysAspAspValGluLeuGluSerValSerGluLysGlyA 583
Db 13884 --CGTACCGCTGTGC-----GCGC 13866
QY 583 spSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSerAnThrA 603
Db 13865 ACCAGCGCGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13815
QY 603 rgSerGluArgGlnLysLysArgMetTyrTyrGlyHisSerLeuGluGluAspLeuG 623
Db 13814 GCGCGCTGCGCAGCG 13755
QY 623 luTrpSerGluPro----- 627
Db 13754 ATCAGCGCAGCGCGCGCTCGGGTCTCGCTTCGCGCGCGCGCGCGCGCGCGCGCG 13695
QY 628 -----GlnLeuLysAspSerGlyValAspThrCysSerSerThrThrLeuAeng 644
Db 13694 GCG 13635
QY 644 luGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLysAspGlyAspA 664
Db 13634 ATCCACACAGCTCGGTTCGAGCGCGCGCTCCGACAGCGCGCGCGCGCGCGCGCG 13575
QY 664 rgLeuLeuArgLysLeuLeuAsnLys-----ArgLeuLysA 677
Db 13574 AGCGCTGCTGCTCGCTCATGCGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13515
QY 677 spGlySerValPro-----ArgAspSerGlyAlaMetLeuGlyLeuLys----- 691
Db 13514 GCGGTGACGTGACACACACCGCTCAGCGCGCTCGCGCGCTGCTCGAGCGCGCAAC 13455
QY 692 -----ValValGlyLysMetThrGluSerGly--ArgLeuCysAlaPheIleTh 708
Db 13454 AATGATCGAGATCGGCAATCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13395
QY 708 rLysValLysLys----- 712
Db 13394 TTCGTGACCATGACACACAGCGAGCGCGCTCACCAGGAGGTGCTGCGAGGTCCACA 13335
QY 713 -----GlySerLeuAlaAspThr---ValGlyHisLeuArgProGlyAspGluValLeuG 730
Db 13334 GCGTGAAGCTCACCTGATACATATGTTGGCCACACAGCT-----GAACGCGGTC 13284
QY 730 uTrpAnGlyArgLeuGlnGlyAlaThrPheGlu----- 743
Db 13283 CCGAGCTGCTCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13224
QY 744 -ValTyrAnIleLeuLeuGluSerLysProGluProGlnValGluLeuValValSerAr 763
Db 13223 CCGGTGCGCTGCGAGCTGACGAGCGAAGCGCTGGCTGCTGCTGCTGCTGCTGCTGCT 13164
QY 763 gProIleGlyAspIleProArgIleProAspSer-----ThrHisAl 777
Db 13163 TCCACTCGGATCCGAAGCTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13104
QY 777 aGlnLeuGluSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSe 797

Db 13103 AACGGCAACACACGCGCTTGGCCCTTGACATCTGAGAGAAGCTGCGACCATCGTGTGG 13044
QY 797 rValThrSerProMetSerProGlyMetLeuArgAspValProGlnPhe-LeuSer---- 815
Db 13043 AGCGCGGCTCATCAGCGCG-----GATGACCCCATAGCTCGCGCTGCGCGCTG 12996
QY 815 ----- 815
Db 12995 CCCCTCGGCTCCCGGCGCAACACACCGCGCGTAAAGCTTCGCGCTCGACGCGCAGAGC 12936
QY 815 ----- 815
Db 12935 TCCACGAGCGCTCGAACAACACCGGACCTAGTCGAGCGCGCGCTCACGAGCGCTGTAA 12876
QY 816 -----GlyGlnLeuSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValT 834
Db 12875 AACCGGACAGGTCCACTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12816
QY 834 hrIleLeuGlyAlaLysAspLeu-----ProSerA 844
Db 12815 TCGCGCGGCG 12756
QY 844 rgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuProAspArgSerA 864
Db 12755 GCGTCTCCGCGCGCTGCTGGAATGCTGTACAGCGCTACTCTCGCGCG----- 12706
QY 864 spLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys----- 880
Db 12705 -----GCCCGCGCGCTCGCGCTCCGACCGACAGCTGACGCGCGCGCGCGCTCC 12654
QY 881 -----TrpAnGlnThrPheIleTyrSerProValHisArgA 893
Db 12653 TCGCGCGCAGCACACGCGCTCGCGCGAGAACACAGCTCCGACAGCTCCGCGCGCGCG 12594
QY 893 rgGluPheArgGluArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValArgGluG 913
Db 12593 CG 12549
QY 913 luGluSerGluPheLeuGlyLysLeuLeuIleGluLeuThrAlaLeuLeuAspAspG 933
Db 12548 AC-----CACCTTGCACAACTGATGATCCCGA 12519
QY 933 luProHisTrpTyrLysLeuGlnThrHisAspValSerSerLeuProLeu----- 949
Db 12518 AGCCACGATGCTCCG 12459
QY 949 ----- 949
Db 12458 AGCTTCG 12399
QY 950 -----ProArgProSerProTyrLeuPro-----ArgArgG 960
Db 12398 TCGCGCGCGCGCGCTCGCGCGCTCCAGCGAGTAGCGTCCGCTGGAAGCGGTAGTCCGC 12339
QY 960 lnLeuHisGlyGluSerProThrArgArgLeuGlnA-rgSerLysArgLysSerAspSerG 980
Db 12338 AGCTCCACG 12291
QY 980 luValSerAspTyrAspCysGluAspGlyValGlyValValSer-----AspTyrArgH 998
Db 12290 ACTCGTGCCTTCGACGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12231
QY 998 isAnGlyArgAspLeuGlnSerThrLeuSerValPro-----GluGlnValMets 1016
Db 12230 GCTCTCGCGCGCGCTCGCCACAACGTCGCTCGCTCCCTCGCGCGAGGACCGCGCG 12171
QY 1016 erSerAsnHisCysSerProSerGlySerProHis-ArgValAspValIleGlyArgThr 1035
Db 12170 CCAGCGCGCACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12135
QY 1036 ArgSerTrpSerProSerAlaProProProGlnArgAsnValGluGln---GlyHisArg 1054


```
Db 3314 GCACAGTCTCCCGATGAGCTTCGGGATGGCGTTC-----AAAAGTCGGCGACGCTAA 3261
Qy 200 InGlnProAspGlnLysValProArgGlyLeuArgAsnGlu-----GluAlaP 216
Db 3260 AGGTGTATTGCGGGGGCGACGTGCATTGCGCAGTACGCTGATGACATCGAGGAA 3201
Qy 216 roGlnGlnLysLysAlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspL 236
Db 3200 CGTCCGAACCTCGCGTGAAGATACTC-----GGTCTCGCGAGC 3162
Qy 236 euSerValProAlaValGlu--LysGlyArgAlaHisGlyLeuThrArgGlnAspThrI 255
Db 3161 ATACAGTCCCTCGAGAGAAATCCAGGAAACGTCACGGATCAACTAGCAGA----- 3110
Qy 255 elySAsnGlySerGlyValLysHisGlnIleAlaSerAspMetProSerAspArgLysAr 275
Db 3109 -----TCTGATATCAACGACGAGAGAC 3084
Qy 275 gSer-----ProSerValSerArgAsp----- 282
Db 3083 ATCTCTTCACAGACCGCTGATCATCAATAAGTCGTACTCTCGTCACTTCGAGAAACGG 3024
Qy 283 -----GlnAsnArgArgTyrGlnGlnSerGluGluArgGluAsp----- 295
Db 3023 ATGCTTTTGACTCCGCGCATCGAGATGTCGAAGCGGAATACGACGAATCGATTCAGT 2964
Qy 296 -----TyrSerGlnTyrValProSerAspGlyThrMetPr 307
Db 2963 ACACGGGGCGTGGCGCGGTACATATTCGTGGACATTCACATCGCGCGGACAGCAAC 2904
Qy 307 oArgSerProSerAspTyrAla---AspArgSerGlnArgGluProGlnPhe----- 324
Db 2903 AAGGAGCGCGAATAATCATGTTTCATCATAGACGATCTCTGTGTAGCCGTTTCTCTCT 2844
Qy 325 -----TyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGlyHisAr 342
Db 2843 TGTGTGATATCTCCAGAC-----GATGGCTCTCGACGAGATGTACG 2802
Qy 342 gHisSerLysGluTyrIleValAspAsp--GluAspVal---GluSerArgAspGluTyrG 361
Db 2801 ACACCAA-----ATCGATTTCGAGACACATCAAGGAGAACCACTCTCTGTACG 2754
Qy 361 lu---ArgGlnArgArgGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuA 380
Db 2753 AAGTTGTGAACGATCCGACGAGAGTACCAGCAGCGCAAGCA-----CTCG 2703
Qy 380 laArgTyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValS 400
Db 2702 AATCGCAACTGGAGATGGCGGAGCAGGTCCAGAAACAGATGCGTGACCGCATGGAGGTCA 2643
Qy 400 erArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuA 420
Db 2642 AGAATAATAGTATCCCGACGAAA-----ACGACGACGGTCCGACGACCGACTTCCCA 2589
Qy 420 spSerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluA 440
Db 2588 GCTCAGCGGCTGCTCGACCAATTCACACGCTCGTGGAGTCTCGCGACATCGAAGA 2529
Qy 440 rgArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgTyrArgGluAlaGlnG 460
Db 2528 GGAGGGCGGTCAACGACACAGTACAGCCGTATCGACCGGGAAACGCCGAATTGATTA 2469
Qy 460 lyGlnSerSerTyrProGln-----ArgThrSerAsnH 471
Db 2468 CGACTACGAGTTTCCATCGGGTGGACGAGCGGAGCGGTCTCTTACGACGAGGAACC 2409
Qy 471 isSerPro-----ProThrProArgArgSerProIleProLeuAspArgProAspMetA 489
Db 2408 ATCGTCCATTCGGTCCCGTCCGAAACAGCGATCGGTTCAGCAAGAGCA-----C 2358
Qy 489 rgArgAlaAspSerLeuArgLysGlnHisHisLeuAspProSerSerAlaValArgLysT 509
```


Qy	357	gAspGluTyrGluArgGlnArgGluGluGluTyrGlnAlaArgTyrArgSerAspPr	377
Db	21917	AGGGGGGGTTCCTCGCAGGCCATCGAGAGGGGGAATATCGACGGCTCGACGGCGCAGCTCC	21976
Qy	377	o-----AsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGlnMe	393
Db	21977	ACGTGACGGCGACGACGAGCGCGCGCCCTTGC---CTGTCTCT---TCCACCCCTCG	22030
Qy	393	tArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSerLe	413
Db	22031	CGAGCTTTCGCCACGAGCGCAAGACGAGCGACGCTCGAGC-----CT	22075
Qy	413	uAla-----AsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspAr	430
Db	22076	GGGGCTACCGATCATCGTGAAGCCTCTGACCACGC-----CACCAAGCCCGCGCG	22126
Qy	430	gProSerArgGlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTy	450
Db	22127	ACTTGGCCGGCAC-----CTGGCTCTCTGTGTCCCGCCGCTCTGCACGACGACGCGC	22180
Qy	450	rSerMetGluArgThrArgGluAlaGln---GlyGlnSerSerTyProGlnArgThrSe	469
Db	22181	TCGCTCCGGCTCACGAGGCGCTCGCGCGCGCGCGCGCTCTCTCGCGTGGCGCC	22240
Qy	469	rAsnHisSerProProThrPro-----ArgArgSerProIlePro-----Le	483
Db	22241	TGAGCCAGGCGCCACCTGGAACCGGAGGCTCTCGCCGAGCACCTCGCCAGGCTTCGCGCG	22300
Qy	483	uAspArgProAspMetArgArgAlaAspSerLeuArg-----LysG	497
Db	22301	AGACCGCGCGCTCGGGCGGTCTCTGCTCTCTGCTCTGACGAAAGTCCCTCGCGCG	22360
Qy	496	-----LysG-----LysG	497
Db	22361	ACCATGCGCGGTGCGCGGACTCGCTCTCTGCTCACCTCGTCCAGCGCTCGCGCG	22420
Qy	497	nHisHis-----LeuAspProSerSerAlaValArgLysThr--	509
Db	22421	ACATCGCTCTGACGCGCCTTGTGTGCTCTTACCGCGCGCGCTCTCTCTCGGACACT	22480
Qy	509	-----LysG-----LysG	509
Db	22481	CCGACCCCATCGCCATCGACGCGCGATGATCGTGGGCTGGCGCGCTGTCGCGCC	22540
Qy	510	-----LysArgGluLysMetGluThrMetLeuAr	519
Db	22541	TCGAGCACCCGAGCGCTGGGAGGGCTGTCGAGTGGCGCGCAGCATCGACGCGAGCG	22600
Qy	519	gAsnAspSerLeuSerSerAspGlnSerGluSerValArg-----Pr	533
Db	22601	CGTGGGCGCTGTCTCCCGTCTCGCCCTCGGACGATGAGGACGAGCTGCTCTCC	22660
Qy	533	oPro-----ProProArgProHisLysSerLysGlyGlyLysMe	547
Db	22661	GCCCGCGCGGTCTACGCTCGCGCTGCTCGCGCTCGGCTCGG-----CGAGC	22711
Qy	547	tArgGlnValSerLeuSerSerGluGluLeuAlaSerThrProGluTyrThr--	566
Db	22712	-CGCGCGCGCAGTACCTTCAAGCCCGGAGCACCTCTCATCACGCGGACCGCGC	22770
Qy	567	-----SerCysAspAspValGluLeuGluSerGluSe	577
Db	22771	GCCGTGGCGCTCAGTTCGCCGATGCTGCTCGAGAGGGCGCAGACCTGTCTCTC	22830
Qy	577	rValSerGluLysGlyAspSerGlnLysArgLysThr-----	591
Db	22831	ATCAGCGCGAGGGGCCAGGCGCGCTCGAGCTCCACGCCGAGTCAAGCGC	22890
Qy	592	-----SerGluGlnGlyValLeuSerAspSe	600
Db	22891	CTGGGCGCGCGTCACTTCCTCGCGCGGTGTGATGTGCGCGACAGAGCGCTGTGCGCAGC	22950

QY	600	rAsnThrArgSerGluArgGlnLysLysArg	---MetTyrTy	613
DB	22951	CTTCTCGAGCAGCTCGACGCCGGAAGGTCGCGATCCGCGCCGCTGTTCACACGCGCGCGC	23010	
QY	613	rGlyGlyHisSerLeuGluGlu-----AspLeuGluTTrpSerGluProGlnLysLys--	630	
DB	23011	ATCGGGCGCCACGCTCCGCTCCGCGCCACCTCTCTCATGAGCTCCGCAAGTGTGTCTCT	23070	
QY	631	-----AspSerGlyValAspThrCysSer--SerThrLeuAsnGluG	645	
DB	23071	GCCAAAGTCTTAGCGCAGGAAACCTCCACAGCACTGCTCGGTCTCGACCCCTCGACGC-	23129	
QY	645	luHisSerHisSerAspLysHisProValThrTTrpGlnProSerLysAspGlyAspArgL	665	
DB	23130	-----CTTCGTCTCTTTCTCGTCCATCGCAGCGCTCGGGCGCGCGA	23172	
QY	665	eulleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAspSerG	685	
DB	23173	CAACAAGCCGGATACGCCGCCGGAACGCCCTT-----CCTCGACGCCCTG	23217	
QY	685	lyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSerGlyArgLeu---	703	
DB	23218	GCCGACCAAGCGCGCAGTCTTGACAGCCGACAGCTCGGTGTGGGGCGGTGGGC	23277	
QY	704	-----CysAlaPheIleThrLysValL	711	
DB	23278	GCGCGCGTGGTATATTACCGCGGCCCTTCGCGAGCCAGCTGAGCAACGTCGTCTGTCG	23337	
QY	711	ysLysGlySerLeuAlaAspThrValGly--HisLeuArgProGly-----AspG	727	
DB	23338	CGATGGCCCTTCGCTCGCCGTGGCGCGCTCGCGCAAGCCCTGAGACACGACGAGACC	23397	
QY	727	luValLeuGluTTrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnI	747	
DB	23398	ACCGTACCGTCGCGCAGATCATGACTGGGCGCGCTTTCGCGCTTCGATCAGCTCGCTCGC	23457	
QY	747	le-----IleLeuGluSerLysProGluProGlnVal-----GluLeuValV	761	
DB	23458	TCCCGCGCTCTCGCGCACTTTCGCGAGCGCGCCCTCGAAGACAGAGAAGCGCGT	23517	
QY	761	alSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGluS	781	
DB	23518	CCTCTCCGAGCACCGCCCGGCCCGCCACCTCTCG-ACAACTCCGAGCGCGCTCGGA	23576	
QY	781	erSerSerSerPheGluSerGlnLys-----MetAspArgProSerIleSerV	798	
DB	23577	GAGCGAGCAGCTCCGCTCTCTCGCGCGCTGTGTGCGAGACAGCGCCCTCTCTCTCGG	23636	
QY	798	alThrSerProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnL	818	
DB	23637	CCACGAAGCGCCCTTCCCAG-----CTCGACCCGACAAAGGCTTCTTCGACCTCGG	23687	
QY	818	eusLysLysLeuTTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyA	838	
DB	23688	TCTCGATTCTGA-----TCATGACCGTTCGAGCTTCGTCTCGCGC	23723	
QY	838	laLysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrP	858	
DB	23724	CTTGCAACAGGCAACCGGCATCAGCTCCGCGCCACCTTCGCTTCGACCATCCCTCTCC	23783	
QY	858	heLeuProAspArgSerAsp-----	864	
DB	23784	TCATCGCTCGCGCTCTTCATCGCGCACTCGCTCGCCACGCGCTCGGCACGAGGCTCTC	23843	
QY	865	-----LysAsnLysArgArgThrIleThrValLysLysThrLeuGluProLysTTrpAsnG	883	
DB	23844	CGCCGAGCGCAGCGCGCGCGCTCCGCGCGCGCTCGAGCGACGACGCCCATCCCATCGT	23903	
QY	883	lnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeuGluIleR	903	
DB	23904	CGGCATGCG--CCTCGCGCTCGCGGCGCGCTCGCGGATGTGACGCTCTTCTGGAGATT	23961	
QY	903	hrLeuTTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGluIleLeu-	922	

QY 219 sLysAlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValPr 239
Db 21216 |||||:::CATGGCGCGGTGGAGTGC CGGAGCGCGAGGTGCAGCGCGCTCC 21262
QY 239 oAlaValGluLysGlyArgAlaHisGlyLeuThrArgGln--AspThrI1 255
Db 21263 AGCGCTATGGCGATCGGCTCTCATCGGGCGATCAACAGCCCTCGTTTCACAGCATCT 21322
QY 255 eLysHengly-----SerGlyValLysHisGlnIle-- 265
Db 21323 CGGGGAGCCCGCTGCGCTGCGCGCCCTGCTCGCGATCTGGAGTCGAGGGGCTTTCG 21382
QY 266 -----AlaSerAspMetProSerAspAr 273
Db 21383 CCTCAAGCTGAGTTACGACTTCGCTCCCACTCCGCGCAGGTGAGTGCATTCGCGAGC 21442
QY 273 g---LysArgSerProSerValSerArgasp-----G1 283
Db 21443 AGCTCTCGATCTCTGTGTGCTCGAGCGCGCTCGACGGCGGTCCCGTTCCTACTCCA 21502
QY 283 nAsnArgArgTyrGluGlnSerGluArgGluAspTyrSerGlnTyrValProSerAs 303
Db 21503 CGGTGAGCGCGCGCGATCGACGGGAGCGAGTTCGACGCGGCTACTGTACCGGAACC 21562
QY 303 p----- 303
Db 21563 TCCGGCAGCGGTTCGCTTCGACAGCGTGTGCAAGCGCTCTTTCGCGGAGAACATCGCT 21622
QY 304 -----GlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArgI 321
Db 21623 TCTTCGTGAGGTGAGCCCGAGTCTGTGTGCTGACCTTGGCTTCGACGAGCTCTCGAAG 21682
QY 321 uProGlnPheTyrGluGlnProGlyHisLeuAsnTyrArgAspSerAsnArgArgGlyHi 341
Db 21683 CGTCGGAGCGCTCGCGCGGTGTGTGCTCTCTGTG-----GAGCGACGAGGGGATC 21736
QY 341 s-----ArgHisSerLysGluTyrIleValAs 350
Db 21737 TAGCGCGTTCCTCGTCTCGCTCTCGAGCTCTACGTCAACGCTTCGCGCTCGATTGGA 21796
QY 350 pAspGluAspVal----- 354
Db 21797 CGACGATCTGCCCCCGGGAAGGGGTGCGCTGCCCTACCTACCCCTTCCAGCGGAGC 21856
QY 355 -----GluSerAr 357
Db 21857 GCTTCTGGCTCGAGCGCTCCACGCGACCGCGCGCGGTCAACACCTTGTCTCCGCTCG 21916
QY 357 gAspGluTyrGluArgGlnArgGluGluGluTyrGlnAlaArgTyrArgSerAspPr 377
Db 21917 AGGGCGGTTCGCGAGCGCATCGAGAGCGGGAATATCGACGCGCTCGAGCGCGAGTCC 21976
QY 377 o-----AsnLeuAlaArgTyrProValLysProGlnProTyrGluGlnIle 393
Db 21977 ACCTGAGCGCGGAGGAGCGCGCGCGCTTCG---CTGCTCCT---TCCACCCCTCG 22030
QY 393 tArgIleHisAlaGluValSerArgAlaArgHisGluArgHisSerAspValSerLe 413
Db 22031 CGAGCTTCGCCACGAGCGGCAAGAGCGGCGAGCGGTGCGACGC-----CT 22075
QY 413 uAla-----AsnAlaGluLeuAspSerArgIleSerLeuLeuArgMetAspAr 430
Db 22076 GCGCGTACCCGATCACGTGGAAGCTCTGACACCGC-----CACACCGCGCGCG 22126
QY 430 qProSerArgGlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTy 450
Db 22127 ACCTGGCGGCAC-----CTGGCTCTCTGTCGCGCGCGCTCTGAGAGGAGCGCGC 22180
QY 450 rSerMetGluArgThrArgGluAlaGln---GlyGlnSerSerTyrProGlnArgThrSe 469
Db 22181 TCCCTCGCGCTCACGAGCGCTCGCCCGCGCGCGCGGTCTCTCGCGGTGGCGC 22240
* Db 22181 TCCCTCGCGCTCACGAGCGCTCGCCCGCGCGCGCGGTCTCTCGCGGTGGCGC 22240

469 rAsnHisSerProThrPro-----ArgArgSerProIlePro-----Le 483
22241 TGAGCAGCGCCACCTGAGCGCGAGGCTCTCGCGAGCACCTGCGCAGGTTGCGCGC 22300
QY 483 uAspArgProAspMetArgAlaArgAlaAspSerLeuArg----- 495
22301 AGACCGCGCGCTCGCGCGGTGCTCTCGCTCTCTCGCTCGACGAAAGTCCCTCGCGC 22360
QY 496 -----LysG1 497
22361 ACCATGCGCGCTGCGCGGAGTCTGCTCTCTCGCTCACCTCTCTCAAGCCCTCGCGC 22420
QY 497 nHisHis-----LeuAspProSerSerAlaValArgLysThr-- 509
22421 ACATCGCTCTGACGCGCCCTTGTGCTCTTTCACCGCGCGCTCTCTCGCTCGACACT 22480
QY 509 ----- 509
22481 CCGACCCCATCGCCCATCGACGCGGATGACTCTGGGCTCGGCGCGCTGCTCGGC 22540
QY 510 -----LysArgLeuLysMetGluThrMetLeuAr 519
22541 TCGACGACCCCGAGCGCTGGGAGGCTCTGTCGAGCTCGCGCGAGCATCGACGCGAGC 22600
QY 519 gAsnAspSerLeuSerSerAspGlnSerGluSerValArg-----Pr 533
22601 CCGTGGCGCTTGTCTCGCGCTCTCGCGCTCGCAACGATGAGACCGAGCTCGCTCTCC 22660
QY 533 oPro-----ProProArgProHisLysSerLysLysGlyGlyLys 547
22661 GCGCGCGCGGTCTACGCTCGCGCTCTGTCGCGCTCGCTCGCTCGCTCGC 22711
QY 547 tArgGlnValSerLeuSerSerSerGluGluGluLeuAlaSerThrProGluTyrThr-- 566
22712 -CGCGCGCGACGATCTCAAGCCCGAGCACCTCTCTCATCGAGCGGACCGCGC 22770
QY 567 -----SerCysAspAspValGluLeuGluSerGluSe 577
22771 GCGCTGCGCTCACTCGCTCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 22830
QY 577 rValSerGluLysGlyAspSerGlnLysGlyLysArgLysThr----- 591
22831 ATCAGCGCGAGGCGCGCGCGCTCGAGCTCGAGCTCGAGCTCGAGCTCGAGCT 22890
QY 592 -----SerGluGlnGlyValLeuSerAspSe 600
22891 CTGGCGCGCGCTCACCTTCGCGCTGTGATGTCGCGAGCGAGCTGTGCGCAGC 22950
QY 600 rAsnThrArgSerGluArgGlnLysLysArg-----MetTyrTy 613
22951 CTCTCGAGCATCGACCGCGAGGTGCGAGGTGCGCGCTGTCACGCGCGCGCGC 23010
QY 613 rGlyGlyHisSerLeuGluGlu-----AspLeuGluTyrSerGluProGlnIleLys-- 630
23011 ATCGGCGCGCATCGCTCGCTCGCGCGCTCTCTCATGAGCTGCGCGAGCTGTCTCT 23070
QY 631 -----AspSerGlyValAspThrCysSer-SerThrThrLeuAsnGluG 645
23071 GCGAAGTCTTAGCGCGAGGAACCTCCACGACCTCTCGCTCGCTCGAGCTCGAGC- 23129
QY 645 LuHisSerHisSerAspLysHisProValThrTyrGlnProSerLysAspGlyAspArgL 665
23130 -----CTGCT 23172
QY 665 euIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerValProArgAspSerG 685
23173 CAACAAGCGGATACGCGCGCGGAAACGCTT-----CTCGAGCGCGCTG 23217
QY 685 lYalaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSerGlyArgLeu---- 703
23218 GCGACCGCGCGAGTCTTGGACGAGCGGACGCTCGCTCGTGTGTGGGCGCGTGGGC 23277
QY 704 -----CysAlaPheIleThrLysValL 711

Db 23278 GCGCGGCGGTATATTCACGGGCGCCCTCGCAGCCAGCTCGAGCAAGCTCGTCTGTCG 23337
Qy 711 yslYsglySerLeuAlaAapThrValgly---HisLeuArgProgly-----AspG 727
Db 23338 CCGATGGCCCTTCGCTGGCGGTGGCGGCTCGCGAAGCCCTGGAGCAACAGAGACC 23397
Qy 727 luValLeuGluTrpAenGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnI 747
Db 23398 ACCGTCACCGTCGCGACATGACTGGCGCGCTTGGCCCTTCGATCAGCTCGCTCGC 23457
Qy 747 le-----lleuGluSerLysProGluProGlnVal-----GluLeuValV 761
Db 23458 TCCCGCGCTCTCGCGCGACTTGGCGGAGCGCGCCCTCGAAGACAGAGAGCGCGCT 23517
Qy 761 alSerArgProIleGlyAspIleProArgIleProAspSerThrHisalaGlnLeuGluS 781
Db 23518 CTCTCTCGAGCAGCGCGCGCCCGCCCGACCTCTCG-ACAAGCTCGGAGCGCGCTCGGA 23576
Qy 781 erSerSerSerPheGluSerGlnLys-----MetAspArgProSerIleSerV 798
Db 23577 GAGCGAGCAGCTCCGCTGCTCGCGCGCTGCTGTGCGAGAGAGCGCCCTCGTCTCGG 23636
Qy 798 alThrSerProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnL 818
Db 23637 CCACGAAGCGCGCTTCCAG-----CTCGACCCGCAAGAGCTTCTCGACCTCGG 23687
Qy 818 euSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyA 838
Db 23688 TCTCGATTTCGA-----TCATGACCGCTCGAGCTTCGTCGCGG 23723
Qy 838 laLysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrP 858
Db 23724 CTTCACAGCAGCCACCGGCATCAAGCTCCCGGCCACCTCGCTCGACCATCCCTCTCC 23783
Qy 858 heLeuProAspArgSerAsp----- 864
Db 23784 TCATCGCTCGCGCTCTTCATGCGCGACTCGCTCGCCCAAGCTCGGCAAGGCTCTC 23843
Qy 865 -----LysAsnLysArgThrLysThrValLysLysThrLeuLupProLysTrpAsnG 883
Db 23844 GCGGAGCGAGCGCGCGCGCTCGCGCGCGCTCGAGCGAGCGAGCCCATCGCATCGT 23903
Qy 883 lnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeuGluIleT 903
Db 23904 CGGCATGCG--CTTCGCGCTCGCGCGCGCTCGCGGATGTCAGCTCTTTGGAGTTC 23961
Qy 903 hrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGluIleLeu- 922
Db 23962 CTCACCAAGGCGCGAGCGCTCGAGCCATTCACAGAGCGCTGGGAGCGCGGTGCC 24021
Qy 923 -----lleuLeuGluThrAlaLeuLeuAsp----- 931
Db 24022 CTCACGACCCCGAGCCCGAGCGCGAGAGCTACGTCGCGCATCGCGCATGCTC 24081
Qy 932 -----AspGluProHisIleTyrLysLeuGlnThrHisAspValSerSerLeuProLeup 950
Db 24082 GACCAGATCGACTTCGACCTTCGACCTTCCTTCGCAATCGCATCGCCCGCGGAGCGCAACAC 24141
Qy 950 roArg-----ProSerProTyrLeuProArgGlnLeuHisGlyGluSer----- 965
Db 24142 CTCGACCCCGAGCAGCGCTGCTCTCGAATCTGCTGCTGGTGGCCCTCGAGAGCGCGGC 24201
Qy 966 -----ProThrArgArgLeuGlnArg-----SerLysA 975
Db 24202 ATCGTCCCGACCTCCCTCAAGGACTCCCTCACCGCGCTCTTCGTCGCGCATCGCGCGCG 24261
Qy 975 rgIleSerAspSer-GluValSerAspTyrAspCysGlu----- 987
Db 24262 GAATACCGGATGCAAGAGCGAGCTCGGAAGTTCCGAGGTTTACTTCATCCAAAGGCAC 24321
Qy 988 -----AspGlyValGly-ValValSerAspTyrArgHisAsnGlyAr 1001
Db 24322 TCCGCGTCTTTGGCGCGCGGCGCTTGGCTATACGCTCGGCTCCAGG-----GGCGG 24375
Qy 1001 gaspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCysSe 1021
Db 24376 CGATCTTCGGTCGACACCGCGCTCTCTCGCTCGCTCTCTCTCGCTCGCTCGCTCGCT 24435
Qy 1021 rProSerGlySerProHisArgValAspValIleGlyArgThr---ArgSerTyrSerPr 1040
Db 24436 GGCCTCCGACAGGCGAGTGCACCTCGCCCTCGCGCGGCGGTGCTCGCTCATGCTCTCC 24495
Qy 1040 oSerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrG 1060
Db 24496 CCCCAGACCTTCGTCATCTTCGCTTCGCGCTTCGCGCGCGCGCGCGCTCCAG 24555
Qy 1060 yHisTyrAsnThrIle-SerArgMetAspArgHisArgValMetAspAspHisTyrSers 1080
Db 24556 ACCTTCTCGGACACCGCGAGCTTACGAGCGGAGAG----- 24593
Qy 1080 erAspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSert 1100
Db 24594 --AGCGTCTGCTCTTCGCTTCGCTTCGAGCGGATCGG----- 24626
Qy 1100 hrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProA 1120
Db 24627 -----CGAGCGCTTCGCGCGGAGAC-ACCGCTCTCTCTCTCTCTCTCTCTCTCT 24680
Qy 1120 spThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerP 1140
Db 24681 CAACACGACGCGCGCT-----CGAGCGTATCACGCGCC 24716
Qy 1140 roAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrP 1160
Db 24717 CAACGCGACCTCCGAGAGAGTCTCTCGCGCGCGCTCCAGCGCGCGCATCA----- 24772
Qy 1160 roGlyThrGlyArgGlyArgGlnLeuProGlnLeuProGlyGlyThrLeuGluA 1180
Db 24773 -----CCCCCGCGAGCTCGAGCTCGT 24794
Qy 1180 rgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnV 1200
Db 24795 CGAGTGCCATG----- 24805
Qy 1200 alAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpA 1220
Db 24806 -----GACCGCGCACCTCGCTGGGAGACCCCATCGAGTGCAG-----CCCTGGC 24851
Qy 1220 spProHisArg--GlyAlaAspThrValSerThrLysSerSerAspSerAspValSerA 1239
Db 24852 CGCGTCTACGCGCGAGCGGAGCCGCTGAAAGCTCTCTCTCTCGCGCGCTCAAGAC 24911
Qy 1239 spValSerAlaValSerArgThrSerSerAlaSerArgPhe----- 1252
Db 24912 CAACATCGGCGCATCTCGAGCGCGCTCGCGCTCGCGCGCTCGCAAGATGGTTCGCGTC 24971
Qy 1253 --SerSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleS 1272
Db 24972 GCTCGCGCAGCGCGCTCGCCCGCCCGCTCCAGCGCGAGCCCGAC-----G 25016
Qy 1272 erValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrL 1292
Db 25017 CAATCCCTCATCG-----AGTGGAGCGCTCGCCAT 25049
Qy 1292 ysserThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerA 1312
Db 25050 CGAGCTGCTGATACCCCGAGGCTT---GGCCC-CGCGAGAGATGGAGTCCCGCGC 25105
Qy 1312 spThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgArgSerSerIleGlyAlaL 1332
Db 25106 GCGCGCGCATCTCGCGCTCGGATTCGCGGACCAAGCGCGCGCTCATCTCGAGAGG 25165
Qy 1332 ysmetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeu----- 1349
Db 25166 CTCCCGCGCGCTGCGCGCGAGCGCGCACCTCACAGCGCGCGCTCGCAAGCGCTCCCG 25225

Qy	1350	-----SerGlnThrGluGlyGlyGlyLeuL	1358
Db	25226	CGCGTGGCCGCTGCTCTGTCCGCCAGGAGCGAGCCCGCTCGCGCGCCAGCGAAGC	25285
Qy	1358	ysLeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnT	1378
Db	25286	GGTCCGCGACCACTGTCGCCACAGACACTCACCTCGCGGATGTGGCTATTGCG	25345
Qy	1378	rpMetThrArgGlnAla-	1387
Db	25346	TGGCCACACCCGCGCCACTTCGAGACACGCGCCGCTCTCGTAGCCCAACACCGCAGC	25405
Qy	1387	erThrAspGlySerMetAsnSerTyrSer-	1396
Db	25406	AGTCTCTCTCGCGCTGACTGCTCGCCCGCTCTTCGCGCTCAGCTCGAAGCAT	25465
Qy	1397	-----SerGluGlyAsnLeu-----IlePheProGlyVal-----	1406
Db	25466	GACGAGCGAAGCCAGCGCAAGCTGCTTCGTCTTCTGGGCAGAGCTCGCAGTGG	25525
Qy	1406	-----	1406
Db	25526	AAGGATGGCCCTCTCGCTGCTGACTCTCGCCCGCTCTTCGCGCTCAGCTCGAAGCAT	25585
Qy	1407	-----ArgLeuAlaSerAspSerGlnPheSer-----	1415
Db	25586	GCGAGCGCGCTCGCTCTCTCAGCTGAGTGGAGCCTGCTCGCGTCTCGCGCGCGAAG	25645
Qy	1416	-----AspPheLeuAsp-----	1419
Db	25646	AGGGCGCCCTCCCTCGACCGGTGACGCTCGTACAGCCCGCTCTTTGCCGTCATGG	25705
Qy	1420	-----GlyLeuGlyProAlaGlnLeuValGlyArgG	1430
Db	25706	TCTCTCTGGCGCCCTCTGGCGCTCGCTCGGCGTACAGCCCGCGCTCGTGGCGCAC-	25764
Qy	1430	lnThrLeuAlaThrProAlaMetGlyAspIleGln-----ValGlyMetMetA	1446
Db	25765	AGTCAGGGCGAGATCGCGCGCTTCGTGCGAGCGCTCTCT	25807
Qy	1446	spLybLybGlyGlnLeuGluValGluIleAlaArgGlyLeuValVallybProG	1466
Db	25808	CCCTCGAGGAGCGGCCCGCATCGCGCTCGGAGCAAGCGCTCACCCGTCGCG	25867
Qy	1466	lySerLybThr-----LeuProAlaProTyrVallybValTyrLeuLeuA	1481
Db	25868	GCAACGGGGCATGGCGCGCGTGCAGCTCGGCGCTCGACCTCCAGACCTACCTCGCTC	25927
Qy	1481	spAsnGlyValCysIleAlaLybLybThrLybValAlaArgLybThrLeu-----	1498
Db	25928	CCTGGGCGACAGGCTCTCCATCGCGCGCTCACAGCCCCAGGCGCAGCTCGTGTCCG	25987
Qy	1499	--GluPro-----LeuTyrGlnGlnLeuLeuSerPheGluGluSerProG	1513
Db	25988	GCGAGCCCGCGCCATCGACGCGCTGATCGCTCACCGCAGCGAG-----	26037
Qy	1513	lnGlyArgValLeuGlnIleValTyrGlyAspTyrGlyArgMetAspHisLybSerP	1533
Db	26038	-----GTCTCGCCGCAAAAGTCCGCGCTGACACTACGCTCC	26074
Qy	1533	heMetGlyValAlaGlnIleLeuLeuAspGluLeuGlu-----LeuSerAnMet-Val	1550
Db	26075	ACTCCGCCAGATGGAGCGCGTCCAGACAGCTCGCGCAGGTTAGCCAAATCGCTC	26134
Qy	1551	IleGlyTrpPheLybLeuPheProProSerSerLeu-----ValAspProThrSerAlaPro	1569
Db	26135	CTCGGA-----CGTGGAGCTCCCTCTTTATTCGACCGCTCACCGCACCA	26179
Qy	1570	LeuThrArgArgAlaSer-----GlnSerSerLeuGluSerSerThrGlyProSer	1586
Db	26180	GGCTCGACGGCTCCGAGCTCGACGGCGCTACTGTGTATCGAACTTCGGGAAACCGTCC	26239

[illegible]

STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/457,646A
 FILING DATE: 01-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/457,205
 FILING DATE: 01-JUN-1995
 APPLICATION NUMBER: 08/258,261
 FILING DATE: 08-Jun-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Elmer, James Scott
 REGISTRATION NUMBER: 36,129
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8614
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28958 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-457-646A-6

Alignment Scores:
 Pred. No.: 1.36e-07 Length: 28958
 Score: 270.50 Matches: 436
 Percent Similarity: 30.17% Conservative: 194
 Best Local Similarity: 20.88% Mismatches: 768
 Query Match: 3.26% Indels: 696
 DB: 1 Gaps: 90

US-09-617-099B-1 (1-1590) x US-08-457-646A-6 (1-28958)

```

QY 132 CysGlyIleCysHisLysThrLysPheAlaAspGlyCysGlyHisLysCysSerTyrCys 151
DB 20882 TGTCTGATACCTCGCGCGTCTTCC-----GGGCACAGCTCGAAGCGTGG 20926

QY 152 Gln-----ThrLysPheCysAlaArgCysGlyGly-----Arg 162
DB 20927 AGCGCGCTCGCGCCCGACGCTGCTCGCGGTCTCGCGGTCTCGCGCGGAGAGG 20986

QY 163 ValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGlnGlu 182
DB 20987 GCGCGCCCGCTCGACCGGCTCGAGTGTGTGCGCGCGCTGTCTCGATGATGCT 21046

QY 183 IleLeuThrLysSerGlyAla-----189
DB 21047 CGCTGCGCGCTGTGCGCTCATGGCGTCCAGCCGCGGTGTCGCGCATAGCC 21106

QY 190 -----TrpPheTyrAsnSerGlySerAsnThrLeu-GlnGlu 201
DB 21107 AGGCGAGATCGCGCGCTGTGTGCGCGCGCTGTGCTCGAGGACGTGCGCAGC 21166

QY 201 nProAspGlnLysValPro-----ArgGlyLeuArgAsnGluAlaProGlnGlu 219
DB 21167 TGTGCGCGCTCGCGCGCTGTGCTGTGAGCTCGCGCGCGCGCGCGCGCGCG 21215

QY 219 sLysAlaLysLeuHisGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValPr 239
DB 21216 -----CATGCGCGGTGAGCTGTGCGAGGCGGAGGTCGCGCGCGCTCC 21262

QY 239 oAlaValGluLysGlyArgAlaHisGlyLeuThrArgGln-----AspThr1 255
DB 21263 AGCGCTATGCGGATCGGCTCTCATCGGCGCATCAACAGCCCTCGTTTCAGACGAT 21322

QY 255 eLysAsnGly-----SerGlyValLysHisGlnIle-- 265
DB 21323 CGCGCGCGCGCTCGCGCTCGCGCTCGCGATCTGAGTCTCGAGGCGCTTCG 21382

QY 266 -----AlaSerAspMetProSerAspAr 273
DB 21383 CCTCAAGCTGAGTTACGACTTCGCTCCACTCCGCGCAGGTGCGAGTTCGCGACG 21442

QY 273 g---LysArgSerProSerValSerArgAsp-----Glu 283
DB 21443 AGCTCCTCATCTCTGTGCTGCGCGCTCGAGCGCGCTCGAGCGGCTCTACTTCA 21502

QY 283 nAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrValProSerAs 303
DB 21503 CGGTGAGCGCGCGCGCTCGAGCGGAGCGAGTTCGACCGCGCTACTGTACCGAACC 21562

QY 303 p-----303
DB 21563 TCCGCGCGCGTCCGCTTCGCGAGCGCTGTGCAAGCGCTCTTCGCGGAGAACATCG 21622

QY 304 -----GlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 321
DB 21623 TCTTCGTGAGGTGAGCGCGCGCTGTGCTGACCTTGGCTTGCACGAGCTCTCGAAG 21682

QY 321 uProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgGlyHis 341
DB 21683 CGTCGAGCGCTCGCGCGCGGTGTGCTCTCTGTG-----GAGCGAGAGGAGATC 21736

QY 341 s-----ArgHisSerLysGluTyrIleValAs 350
DB 21737 TAGCGCGCTTCCTCGTCTCGCTCTCGAGCTCTACGTCAACGCGTTCGCGCTTGA 21796

QY 350 pAspGluAspVal-----354
DB 21797 CGACGATCTGCGCGCGCGGAGCGGTCGCGTCCACCTACCCCTTCAGCGCGAGC 21856

QY 355 -----GluSerAr 357
DB 21857 GCTTCGCGCTCGACCGCTCGAGGACCGCGCGCGCTCAACACCTGCTCGCGCTG 21916

QY 357 gAspGluTyrGluArgGlnArgGluGluTyrGlnAlaArgGluTyrArgSerAspPr 377

```



```

Db 25928 CTGGGGGACAGGCTCTCCATCCGCGCGTCAACAGCCCGAGGCGCCAGCTCGTGTCCG 25987
Qy 1499 --GluPro-----||||||
Db 25988 GCGAGCCCGCCGACGAGCGCTATCGACTCGCTACCGGAGCGGAG-----26037
Qy 1513 InGlyArgValLeuGlnIleValTrpGlyAspTyrGlyArgMetAspHisLysSerP 1533
Db 26038 -----GTCTCGCCGAAAGTCCGCGTGCAGTACCGCTCC 26074
Qy 1533 heMetGlyValAlaGlnIleLeuLeuAspGluLeu-----LeuSerAsnMet-Val 1550
Db 26075 ACTCCGCCAGATGAGCGCGTCCAGAGAGAGTCCGCGAGGTCTAGCAACATCGCTC 26134
Qy 1551 IleGlyTrpPhenylsLeuProProSerSerLeu---ValAspProThrSerAlaPro 1569
Db 26135 CTCGGA-----CGTGGAGCTCTCTTTATTGCGACCGTCAACCGGACCA 26179
Qy 1570 LeuThrArgAlaSer-----GlnSerSerLeuGluSerSerThrGlyProSer 1586
Db 26180 GGCTCGACGGCTCCGAGCTCGAGCGCGTACTGTGTATCGAACTCCGGCAACGCTCC 26239
Qy 1587 TyrSerArgSer 1590
Db 26240 TGTCTCGAGCG 26251

```

RESULT 12

```

US-08-458-076A-6
; Sequence 6, Application US/08458076A
; Patent No. 5698425
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,076A
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CSC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-458-076A-6

```

```

Alignment Scores:
Pred. No.: 1,36e-07 Length: 28958
Score: 270.50 Matches: 436
Percent Similarity: 30.17% Conservative: 194
Best Local Similarity: 20.88% Mismatches: 768
Query Match: 3.26% Indels: 696
DB: 1 Gaps: 90

US-09-617-099B-1 (1-1590) x US-08-458-076A-6 (1-28958)
Qy 3 AlaProLeuGlyProArg-----GlyArgProAlaProThrProAlaAlaSerGlnPro 20
Db 20577 GCTCTGTGCGCCAGAGCGAGCGCGTGGCGCCAGGCGCCAGAGCTCCGCGACCA 20636
Qy 21 ProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysIleLe 40
Db 20637 CTCTCTGCCCCACGAGACCT-----CGCCCTGGCGGATGTAGCCTA 20678
Qy 41 LeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeuLysIle 60
Db 20679 CTCGCT-----20684
Qy 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
Db 20685 ---CGCCACCAACCGCGGCTA-CTTCGAGCACCGTGGCGCTCTCTGTGTCCAGCGCGG 20740
Qy 81 ValAsn-----AsnValLeuGlnProGlnGlnLysGlnPro-----92
Db 20741 AAGAGCTCTCTCCGCGCTCGATTGCTGCGCGCGAGGAGCGCCGCGCGAGCACCGTCG 20800
Qy 93 ---AsnGluLysGluProGlnThrLysLeuHisGlnGlnPheGluMetTyrLysGluGln 111
Db 20801 TCGAACGAGCGGAGGACCGACCGGAGTGTCTGCTCTTC-----20842
Qy 112 ValLysLysMetGlyGluGluSerGlnGlnGlnGlnGlnLysGlyAspAlaProThr 131
Db 20843 -----CTGGGC-----AAGGCTCGAGTGGAGGATGGCCCTCTCC 20881
Qy 132 CysGlyIleCysHisLysThrLysPheAlaAspGlyCysGlyHisAsnCysSerTyrCys 151
Db 20882 TGCTCGATACCTCCGCGGTCTTC-----GGGCACAGCTCGAAGCGTCG 20926
Qy 152 Gln-----ThrLysPheCysAlaArgCysGlyGly-----Arg 162
Db 20927 AGCGCGCCTCGCGCCCGCCCGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 20986
Qy 163 ValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGlnGlu 182
Db 20987 GCGCGCGCGCGCTCGACCGGTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 21046
Qy 183 IleLeuThrLysSerGlyAla-----189
Db 21047 CGCTGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 21106
Qy 190 -----TrpPheTyrAsnSerGlySerAsnThrLeu-GlnG 201
Db 21107 AGGCGAGATCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 21166
Qy 201 nProAspGlnLysValPro-----ArgGlyLeuArgAsnGluAlaProGlnGlu 219
Db 21167 TGTGTGGCGTGGCGAGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 21215
Qy 219 sLysAlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValPr 239

```


QY 711 yslYsglySerLeuAlaAspThrValGly---HisLeuArgProGly-----AspG 727
Db 23338 CCGATGCGCCCTTGGCTGGCGGCGCTGGCGCAAGCCCTCGAGCAGCAGAGACC 23397
QY 727 luValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnI 747
Db 23398 ACCGTACCGTCGCGCAGCATCGACTGGCGGCGCTTTGGCGCTTCGATCAGCGTGGCTCGC 23457
QY 747 le-----IleLeuGluSerLysProGluProGlnVal-----GluLeuValV 761
Db 23458 TCCCGCGCTCTCGCGACTGTCGCGAGCGCGCGCTCGAGCAGCAGAGCGCGCT 23517
QY 761 alSerArgProIleGlyAspIleProAspSerThrHisAlaGlnLeuGluS 781
Db 23518 CTTCTCCGAGCAGCGCGCGCGCGCTCTCTCG-ACAAGCTCCGAGCGCGCTCGGA 23576
QY 781 erSerSerSerPheGluSerGlnLys-----MetAspArgProSerIleSerV 798
Db 23577 GAGCGAGCAGCTCTGCTCTGCGCGCGCTGGTGGCAGCAGCGCGCTCTGCTCG 23636
QY 798 alThrSerProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnL 818
Db 23637 CCAGAGCGCGCTCCAG-----CTCGACCCCGCAGCAGCGCTTCTCGACCTCGG 23687
QY 818 euSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyA 838
Db 23688 TCTCGAATCGA-----TCATGACCTCGAGCTTCTCGCGG 23723
QY 838 laiLysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrP 858
Db 23724 CTTGCAACAGCGCCACCGGCATCAAGCTCCCGGCCACCTCGCCTTCGACCATCTCTCC 23783
QY 858 heLeuProAspArgSerAsp----- 864
Db 23784 TCATCGCGTGGCTCTCTATGCGCGACTGCTCGCCAGCGCTCGGACAGGCTCTC 23843
QY 865 -----LysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLysTrpAsnG 883
Db 23844 CGCGAGCGCAGCGCGCGCGCTCGCGCGCGCTCGAGCAGCGCCATCGCATCGT 23903
QY 883 lnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeuGluIle 903
Db 23904 CGGCATGGC--CCTGCGCCTCGCGCGCGCTCGCGATGTCGAGCTCTTTGGAGTTC 23961
QY 903 hrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGluIleLeu- 922
Db 23962 CTCACCAAGGCGCGAGCGGTGAGCCCATTCACAGAGCGCGCTGGGACGCGCGGTGC 24021
QY 923 -----IleGluLeuGluThrAlaLeuLeuAsp--- 931
Db 24022 CTCTAGACCCCGACCCCGAGCGCGCGAGCTAGCTCGCGCATGCGCGATGCTC 24081
QY 932 -----AspGluProHisTrpTyrLysLeuGlnThrHisAspValSerSerLeuProLeuP 950
Db 24082 GACCATGACCTTTCGACCTCTCTCGACCTCTCTCGGCATCAGCCCGCGGAGCGCAACAC 24141
QY 950 roArg---ProSerProTyrLeuProArgArgGlnLeuHisGlyGluSer----- 965
Db 24142 CTGACCCCGACCGCGCTGCTCTCGAATCTGCTGGCTGGCTCGAGGACGCGCGC 24201
QY 966 -----ProThrArgArgLeuGlnArg-----SerLysA 975
Db 24202 ATGCTCCCACTCCCTCAGGACTCTCTCAGCGGTCTCTCTCGCATCTCGCGCGC 24261
QY 975 rgIleSerAspSer-GluValSerAspTyrAspCysGlu----- 987
Db 24262 GAATACCGCATCGAAGGCGAGCTCGAAGGTTCGAGGTACTTACTTCATCCAGGCAT 24321
QY 988 -----AspGlyValGly-ValValSerAspTyrArgHisAsnGlyAr 1001
Db 24322 TCCGCTCTCTTGGCGGCGGCTTGGCTTATACGCTCGGCTCGCAGG-----GGCG 24375

QY 1001 gAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCysSe 1021
Db 24376 CGATCTTCGTCGACACCGCT 24435
QY 1021 rProSerGlySerProHisArgValAspValIleGlyArgThr---ArgSerTrpSerPr 1040
Db 24436 GCCCTCCGACAGGCGAGTCAACTCGCCCTCGCGCGCGGCGGTCTCTCTCTCTCTCTCTCT 24495
QY 1040 oSerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGl 1060
Db 24496 CCCCAGACCTTCGTCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24555
QY 1060 yHisTyrAsnThrIle-SerArgMetAspArgHisArgValMetAspAspHisTyrSerS 1080
Db 24556 ACCTTCTCGACACAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24593
QY 1080 erAspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerT 1100
Db 24594 --AGCGCT 24626
QY 1100 hrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProA 1120
Db 24627 -----CGACGCTCTCGCGCGAGAC-ACCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24680
QY 1120 spThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerP 1140
Db 24681 CAACACGACGCGCGCT-----CGAGCGGTATCACCGCCCC 24716
QY 1140 roAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrP 1160
Db 24717 CAACGCGACCTCCACGAGAGTCTCTCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 24772
QY 1160 roGlyThrGlyArgGlyArgGlnLeuProGlnLeuProLysGlyThrLeuGluA 1180
Db 24773 -----CCCCCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24794
QY 1180 rgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnV 1200
Db 24795 CGAGTGCATG----- 24805
QY 1200 alAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpA 1220
Db 24806 -----GCACCGGCACTCTCTCTGAGACCCCATCGAGGTGCAAG-----CCCTGGC 24851
QY 1220 spProHisArg---GlyAlaAspThrValSerThrLysSerSerAspSerAspValSerA 1239
Db 24852 CGCGCTCTACCGCGAGCGAGACCGCTGAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24911
QY 1239 spValSerAlaValSerArgThrSerSerAlaSerArgPhe----- 1252
Db 24912 CAACATCGGCGCATCTCGAGCGCGCTCTCGCGCGCTCTCGCGCGCTCTCGCGCGCTCT 24971
QY 1253 --SerSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleS 1272
Db 24972 GCTCGCGCAGCGCT 25016
QY 1272 erValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerLysLysAsnLeuThrL 1292
Db 25017 CAATCCCTCTCATCG-----AGTGGAGGCGCTCTCGCCAT 25049
QY 1292 ysSerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerA 1312
Db 25050 CGACGCTGCTGATACCCCGAGGCTT---GGCGCC-CGCGACGAAGATGCGAGTCTCTCGCGC 25105
QY 1312 spThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgArgSerSerIleGlyAlaL 1332
Db 25106 GCGCGCGCATCTCGCGCTCTCGGATTTCTCGGCGCACCAACCGCCCATCTCTCGAAGAG 25165
QY 1332 yeMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeu----- 1349
Db 25166 CTCCCGCGCGCTCTCGCGCGCGCGCGCATCTCAGCGCGGTCTCGAGCGCGCTCTCTCTCT 25225
QY 1350 -----SerGlnThrGluGlyGlyLysL 1358

QY 533 oPro-----ProProArgProHisLysSerLysLysGlyGlyLys 547
Db 22661 GCCCGCGCGGTTACCGTCCGCGCTCGCGCGCTCCGCTCGG-----CGAGC 22711
QY 547 tArgGlnValSerLeuSerSerGluGluGluLeuAlaSerThrProGluTyrThr-- 566
Db 22712 -CGCGCGCGCAGGTACCTTCAAGCCCGGAGGACCCCTCTCATCCAGCGAGGACCGGC 22770
QY 567 -----SerCysAspAspValGluLeuGluLeuGluSerGluSe 577
Db 22771 GCCGTGCGCTCACGTCCGCGATGGCTCGTCCGAGAAGCGCGAGACACCTCGTCTC 22830
QY 577 rValSerGluLysGlyAspSerGlnLysGlyLysArgLysThr----- 591
Db 22831 ATCCAGCCCGAGGCGGCCAGCGAGGCGCTCGAGCTCCAGCCGAGCTCACGCC 22890
QY 592 -----SerGluGlnGlyValLeuSerAspSe 600
Db 22891 CTGGCGCGCGCTCACCTTCGCGCGGTGATGTGCGCGAGAGGCGCTGTGCCACG 22950
QY 600 rAsnThrArgSerGluArgGlnLysLysArg-----MetTyrTy 613
Db 22951 CTTCGAGCAGCTCGAGCGCGAAGGTCGAGGTCCGAGTCCGCGCGTGTTCACGCGCGCGC 23010
QY 613 rGlyGlyHisSerLeuGlu-----AspLeuGlnTrpSerGluProGlnLys-- 630
Db 23011 ATCGGCGCGCGCTCGCTCGCGCGCCACCTCTCATGAGAGTCCGCGAGCTGTCTCT 23070
QY 631 -----AspSerGlyValAspThrCysSer-SerThrThrLeuAsnGluG 645
Db 23071 GCCAAGGTCCTAGCGCGAGGAACTCCACGACCTGCTCGGTCTCTCGACCCCTCGACG- 23129
QY 645 luHisSerHisSerAspLysHisProValThrTrpGlnProSerLysAspGlyAspArgL 665
Db 23130 -----CTTCGCTCTTTCTCGTCCATCGCAGCGGTCTCGCGCGCGCGA 23172
QY 665 euLleGlyArgLeuLeuAsnLysArgLeuLysAspGlySerValProArgAspSerG 685
Db 23173 CAACAAGCGGATACCGCGCGGAAACGCTT-----CCTGAGCCCGCTG 23217
QY 685 lValMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSerGlyArgLeu---- 703
Db 23218 GCCGACCGCGCGCAGTCTTGACAGCGCGGACAGTCCGTGGTGTGGCGCGCTGGGCG 23277
QY 704 -----CysAlaPheLeuThrLysValL 711
Db 23278 GCGCGCGGTGGTATATTACGCGCGCCCTCGCGAGCCAGCTGGAGCAACGTCTGTCTCG 23337
QY 711 ysLysGlySerLeuAlaAspThrValGly--HisLeuArgProGly-----AspG 727
Db 23338 CCGATGGCCCTTCGCTGGCGGTGGCGGTGGCGAGCCCTCGAGCAGCAGACG 23397
QY 727 luValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnI 747
Db 23398 ACCGTACCGTCGCGACATCGACTGGGCGCGCTTGGCGCTTTCGATCAGCGTCTCGC 23457
QY 747 le-----IleLeuGluSerLysProGluProGlnVal-----GluLeuValV 761
Db 23458 TCCCGCGCTCTCGCGACTTGGCCGAGCGCGCGCTCGAAGACAGAGAAGCGCGT 23517
QY 761 alSerArgProIleGlyAspLeuProArgLeuProAspSerThrHisAlaGlnLeuGluS 781
Db 23518 CTTCTCCGAGCAGCGCGCGCGCCCGCCCGACCTCTCG-ACAAGCTCCGAGCGCGCTCGGA 23576
QY 781 erSerSerSerPheGluSerGlnLys-----MetAspArgProSerIleServ 798
Db 23577 GAGCGAGCAGCTTCTGCTGCTGCGCGGTGTGTGTGCGAGCAGCGCGCTCTCTCTCG 23636
QY 798 alThrSerProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnL 818
Db 23637 CCACGAGCGCGCTTCCAG-----CTCGACCCCGCAAGGCTTCTTCGACCTCGG 23687
QY 818 euSerIleLysLeuTrpPheAspLysValGlyHisGlnLeuLeuValThrIleLeuGlyA 838
Db 23688 TCCTGATTCGA-----TCATGACCGTCCAGCTTCGTTCGCGC 23723
QY 838 lAlysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrP 858
Db 23724 CTTGCAACAGCGCCACCGGCTCAAGCTCCCGCGCCACCTCGCTTCGACCATCCTCTCC 23783
QY 858 heLeuProAspArgSerAsp----- 864
Db 23784 TCATCGGTGCGCTCTTCATGCGCGACTCGCTGCCCCACCGCTCGCGAGGCTCTC 23843
QY 865 -----LysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLysTrpAsnG 883
Db 23844 CGCGAGCGCGCGCGCGCTCGCGCGCGCTCGAGCGAGCCCATCGCATCGT 23903
QY 883 lnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeuGluLeu 903
Db 23904 CGGCATGCG--CCTGCGCTCGCGCGCGCTCGCGCATGTTCGACGCTCTTTGGAGTTC 23961
QY 903 hrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGluLeu- 922
Db 23962 CTCACCAAGCGCGCGCGCTCGAGCCCATTCACAGAGCCGCTCGGACGCGCGTGC 24021
QY 923 -----lLeuLeuGluThrAlaLeuLeuAsp---- 931
Db 24022 CTCTACGACCCCGACCGCGCGCGCGAGCTACGTCGCGCATGCGCGATGCTC 24081
QY 932 -----AspGluProHisTrpTyrLysLeuGlnThrHisAspValSerSerLeuProLeuP 950
Db 24082 GACCAGATGACCTTTCGACCCCTCTTCGCGCATGACGCCCGCGGAGGCCAAACAC 24141
QY 950 roArg-----ProSerProTyrLeuProArgArgGlnLeuHisGlyGluSer- 965
Db 24142 CTGACCCCGCAGCAGCCCGCTCTCTCGAATCTGCTGCTGCGCGCTCGAGGACCGCGC 24201
QY 966 -----ProThrArgArgLeuGlnArg-----SerLysA 975
Db 24202 ATCGTCCCGACCTCCCTCAAGGACTCCCTCACCGCGCTCTTCGTCGCGCATCTGCGCGC 24261
QY 975 rgIleSerAspSer-GluValSerAspTyrAspCysGlu----- 987
Db 24262 GAATACGCGATGCAAGAGCGGAGCTCGAAGGTTCCGAGGTTTACTTTCATCCAGGCACT 24321
QY 988 -----AspGlyValGly-ValValSerAspTyrArgHisAsnGlyAr 1001
Db 24322 TCCCGCTCTTTCGCGCGCGGCTTGCCTATACGTCTCGCGCTCCAGG-----GGCG 24375
QY 1001 gAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCysSe 1021
Db 24376 CGATCTCGGTGACACCGCGCT 24435
QY 1021 rProSerGlySerProHisArgValAspValIleGlyArgThr---ArgSerTrpSerPr 1040
Db 24436 GCCCTCCGACAGCGCGAGTCAACCTCGCCCTCGCGCGCGGTGCTCATGCTCTCTCTCTCT 24495
QY 1040 oSerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGl 1060
Db 24496 CCCAGACCTTCGTCT 24555
QY 1060 yHisTyrAsnThrIle-SerArgMetAspArgHisArgValMetAspAspHisTyrSers 1080
Db 24556 ACCTTTCGAGCAAGCGCGAGCGGTCTCGAGCGGATCG----- 24593
QY 1080 erAspArgAspArgCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSert 1100
Db 24594 --AGCGCTCGTCT 24626
QY 1100 hrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerGluArgProA 1120
Db 24627 -----CGACGCGCTCGCGCGAGAC-ACCAGCTCTCTCTCTCTCTCTCTCTCTCTCT 24680
QY 1120 spThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerP 1140

Db 24681 CAACACGAGCGCGCT-----CGAGCGGTATCACCGCCCC 24716
Qy 1140 roAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrP 1160
Db 24717 CAACGGACCTCCACGACGAGGTCTCCGGCGCGCTCACGACCGCCGCAFC 24772
Qy 1160 roGlyThrGlyArgGlyArgGlnLeuProGlnLeuProProGlySerThrLeuGluA 1180
Db 24773 -----CCCCGCCGACGTCGAGTCGT 24794
Qy 1180 rgSerAlaMetAspIleGluGluArgGlnMetLysLeuAsnLysTyrLysGlnV 1200
Db 24795 CGAGTGCATG----- 24805
Qy 1200 alAlaGlySerAspProArgLeuGlnAspTyrHisSerLysTyrArgSerGlyTrpA 1220
Db 24806 -----GCACGGGACCTCGCTGGAGACCCATCGAGTGCAG-----CCCTGGC 24851
Qy 1220 spProHisArg---GlyAlaAspThrValSerThrLysSerSerAspSerAspValSerA 1239
Db 24852 CGCGTCTACGCGACGACGACCGCTGAAAGCCTCTCTTCTCGCGGCTCAAGAC 24911
Qy 1239 spValSerAlaValSerArgThrSerSerAlaSerArgPhe----- 1252
Db 24912 CAACATCGGCATCTCGAGCGCGCTCGCGGCTCGCGGCTGCCAAGATGCTCGCCTC 24971
Qy 1253 --SerSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleS 1272
Db 24972 GCTCCGACGAGCCCTGCCCCCACCCTCCAGCGACCCAC-----G 25016
Qy 1272 erValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrL 1292
Db 25017 CAATCCCTCATCG-----AGTGGAGGCGCTCGGCAT 25049
Qy 1292 ysSerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerA 1312
Db 25050 CGACGCTCGTACATCCCGAGCGCTT---GCGCC-CCGACGAAGATGCGAGTCCCGCC 25105
Qy 1312 spThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgArgSerSerIleGlyAlaL 1332
Db 25106 GCGCGCGCATCTCGCGCTCGATTCCTCGGCGACCAACGCGCCATCTCTCGAAGAGG 25165
Qy 1332 ysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeu----- 1349
Db 25166 CTCCGCGCGCTCGCGCGAGCGCCGACCTCAGACGCGCGCTCGAAGCGCTCCG 25225
Qy 1350 -----SerGlnThrGluGlyGlyLysL 1358
Db 25226 CGCGTGGCGCGTCTCTCGCGCAGGAGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCG 25285
Qy 1358 ysLeuArgSerThrValGluArgSerThrGluThrGlyLeuAlaValGluMetArgAsnT 1378
Db 25286 GGTCTCCGACACCTCTCGTCCGCCACGACGACCTCACCCCTCGCGGATGTGGCTATTTCG 25345
Qy 1378 rpMetThrArgGlnAla-----SerArgGluS 1387
Db 25346 TGGCCACACCGCGGCCCATCTTCAGACACGCGCGCTCTCTGTAGGCCACACCGCGAGC 25405
Qy 1387 erThrAspGlySerMetAsnSerTyrSer----- 1396
Db 25406 AGCTCTCTCGCGCTCGATCGTCTCGCCGAGGACAGCGCGCGCGCGCGCGCGCGCGCG 25465
Qy 1397 -----SerGluGlyAsnLeu-----IlePheProGlyVal----- 1406
Db 25466 GACGAGCGGAAGCACGGAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25525
Qy 1406 ----- 1406
Db 25526 AAGGATGGCCCTCTCGCTGCTGACCTCTCGCCCGCTCTCTCGCGCTCAGCTGAGCAT 25585
Qy 1407 -----ArgLeuAlaSerAspSerGlnPheSer----- 1415
Db 25586 GCGAGCGCGCTCGCT 25645

Qy 1416 -----AspPheLeuAsp----- 1419
Db 25646 AGGGCGCCCTCTCTCTGACCGCGCTGACGTCTACAGCCCGCCCTCTTTTGGCGTATGG 25705
Qy 1420 -----GlyLeuGlyProAlaGlnLeuValGlyArgG 1430
Db 25706 TCTCCCTGGCGCGCTCTGGCGCTCTGGCGCTCTGGCGCTCTGGCGCTCTGGCGCTCT 25764
Qy 1430 lnThrLeuAlaThrProAlaMetGlyAspIleGln-----ValGlyMetMetA 1446
Db 25765 -----AGTACGGGAGATCCCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25807
Qy 1446 spLysLysGlyGlnLeuGluValGluIleLeuArgAlaArgGlyLeuValValLysProG 1466
Db 25808 CCTTCGAGGACGGCGCGCGCTCTGGCGCTCTGGCGCTCTGGCGCTCTGGCGCTCTGGCG 25867
Qy 1466 lySerLysThr-----LeuProAlaProTyrValLysValLysValLysValLysLeuA 1481
Db 25868 GCAACGGGCGCATGGCGCGCTCGAGCTCGCGCGCTCCGACCTCCAGACCTACTCTCGCTC 25927
Qy 1481 spAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeu----- 1498
Db 25928 CCTGGGCGACAGGCTCTCCATCGCGCGCTCAACAGCCCGCGCGCGCTCTCTCTCTCT 25987
Qy 1499 --GluPro-----LeuTyrGlnGlnLeuLeuSerPheGluGluSerProG 1513
Db 25988 GCGAGCGCGCGCGCATCGAGCGCTCATCGCTCTACCGCGCGCGAG----- 26037
Qy 1513 lnGlyArgValLeuGlnIleIleValTrpGlyAspTyrGlyArgMetAspHisLysSerP 1533
Db 26038 -----GTCTCGCCGAGAAAGTCCGCGCTCGACTACGCTCTCC 26074
Qy 1533 heMetGlyValAlaGlnIleLeuLeuAspGluLeu-----LeuSerAsnMet-Val 1550
Db 26075 ACTCCGCCAGATGAGCGCGCTCCAGAGAGAGCTCGCGCGAGCTTAGCCACATCGCTC 26134
Qy 1551 IleGlyTrpPheLysLeuPheProProSerSerLeu---ValAspProThrSerAlaPro 1569
Db 26135 CTGCGA-----CGTGGAGTCTCTCTTTATTTCGACCGCTCACCGGCACCA 26179
Qy 1570 LeuThrArgArgAlaSer-----GlnSerSerLeuGluSerSerThrGlyProSer 1586
Db 26180 GGTCTGAGCGCTCGAGCTCGAGCGCGCTCTGTTATCGAAACCTTCGCGCAACCGTCC 26239
Qy 1587 TyrSerArgSer 1590
Db 26240 TGTCTCGAGCG 26251

RESULT 14
US-08-457-335A-6
; Sequence 6, Application US/08457335A
; Patent No. 5723759
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hamner, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,335A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-335A-6

```

Alignment Scores:		
Pred. No.:	1.36e-07	Length:
Score:	270.50	Matches:
Percent Similarity:	30.17%	Conservative:
Best Local Similarity:	20.88%	Mismatches:
Query Match:	3.26%	Indels:
DB:	1	Gaps:
		28958
		436
		194
		768
		696
		90
US-09-617-099B-1 (1-1590) x US-08-457-335A-6 (1-28958)		

Db 24022 CTCTACGACCCCGACCCGACGCGGACGCAAGAGCTACGTCGGCATGCCGCGATGCTC 24081
Qy 932 -----AspGluProHisTyrPtyrLysLeuGlnThrHisAspValSerSerLeuProLeuP 950
Db 24082 GACCAGATCGACCTCTTCGACCCCTGCTTCTTCGGCATCAGCCCGGAGGCGCAACAC 24141
Qy 950 roArg---ProSerProTyrLeuProArgArgGlnLeuHisGlyGluSer----- 965
Db 24142 CTGACCCCGACGACCCGCTGCTCTCGAATCTGCTGGCTGGCTGGCCCTCGAGAGCGCGCG 24201
Qy 966 -----ProThrArgArgLeuGlnArg-----SerLysA 975
Db 24202 ATCGTCCCGACCTCCCTCAAGACTCCCTCACCGGCTCTTCGTCGGCATCTCGCGCGC 24261
Qy 975 rgfLeSerAspSer-GluValSerAspTyrAspCysGlu----- 987
Db 24262 GAATACGCGATGCAAGAGCGAGCTCGAAGGTTCGAGGTTTACTTCACCAAGCACT 24321
Qy 988 -----AspGlyValGly-ValValSerAspTyrArgHisAsnGlyVal 1001
Db 24322 TCGCGCTCTTTGGCGGGGGCTTGGCGCTATACGCTCGGGCTCCAGG-----GGCGG 24375
Qy 1001 gAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCysSe 1021
Db 24376 CGATCTTCGTCGACACCGCT 24435
Qy 1021 rProSerGlySerProHisArgValAspValIleGlyArgThr---ArgSerTyrSerPr 1040
Db 24436 GCGCTCCGACAGGCGAGTGCACCTCGCGCTCGCGCGCGGCTGCTGCTCATGCTCTCC 24495
Qy 1040 oSerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGl 1060
Db 24496 CCGACACTTCGTCTATCT 24555
Qy 1060 yHisTyrAsnThrIle-SerArgMetAspArgHisArgValMetAspAspHisTyrSerS 1080
Db 24556 ACCTTCTCGACACCGCGCTACGAGCGGAGA----- 24593
Qy 1080 eAspArgAspArgAspCysGluAlaAspArgGlnProTyrHisArgSerArgSerT 1100
Db 24594 --AGGCGCTGCTGCT 24626
Qy 1100 hrGluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProA 1120
Db 24627 -----CGAGCCCTCCCGCGAGAC--ACCGGCTCTCTCTCTCTCTCTCTCTCTCTCT 24680
Qy 1120 spThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerP 1140
Db 24681 CAACACGACGCGCGT-----CGAGCGGTATCACCGCCCT 24716
Qy 1140 roAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrP 1160
Db 24717 CAACGCACTCCCGACAGAGTCTCGCGCGCGCTCCACGACGCGCGCATCA---- 24772
Qy 1160 roGlyThrGlyArgArgGlnLeuProGlnLeuProProLysGlyThrLeuGluA 1180
Db 24773 -----CCCCCGCGAGCTCGAGCTGT 24794
Qy 1180 rgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnV 1200
Db 24795 CGAGTCCCATG----- 24805
Qy 1200 alAlaGlySerAspProArgLeuGlnGlnAspTyrHisSerLysTyrArgSerGlyTTPA 1220
Db 24806 -----GCACCGGCACCTCGCTGGAGACCCCATCGAGGTGCAAG-----CCCTGGC 24851
Qy 1220 spProHisArg---GlyAlaAspThrValSerThrLysSerSerAspValSerA 1239
Db 24852 CGCGGTCTACGCGACGAGAGCCCGCTGAAAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24911
Qy 1239 spValSerAlaValSerArgThrSerSerAlaSerArgPhe----- 1252
Db 24912 CAACATCGGCGCATCTCAGCGCGCTCTCGGCGCTCGCGGCGTCCGCAAGATGTCGCTC 24971

Qy 1253 --SerSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsnArgLysIleS 1272
Db 24972 GCTCCGCGACGAGCCCTGCCCCCACCCTCCACGCGACCCAC-----G 25016
Qy 1272 erValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrL 1292
Db 25017 CAATCCCTCATCG-----AGTGGAGCGCTCGCCAT 25049
Qy 1292 ysSerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerA 1312
Db 25050 CGACGCTGCTGATACCCGAGGCTT---GGCCC--CGCCACGAAGATGCGATCCCCGCC 25105
Qy 1312 spThrAlaValGlyAlaLeuGlyThrSerGlyLysArgArgSerSerIleGlyAlaL 1332
Db 25106 GCGCGCGCATCTCCGCTTTCGATTCGCGGACCAACGCGCCATCTCTCTCGAAGAG 25165
Qy 1332 ysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeu----- 1349
Db 25166 CTCCCGCGCTCTCGCGCGAGCGCCACCTCACGCGGCTCGAAGCGCTCCG 25225
Qy 1350 -----SerGlnThrGluGlyGlyLysL 1358
Db 25226 CGCGCTGCGCGCTGCTCTGTCGCGACGAGCGCGCTGCGCGCCAGCGAAGC 25285
Qy 1358 ysLeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnT 1378
Db 25286 GCGCTCGGACACCTCGTCCGCCACGACGACCTCACCTCGCGATGCGCTATTGCG 25345
Qy 1378 rpMetThrArgGlnAla-----SerArgGluS 1387
Db 25346 TCGCCACACCGCGCGCCACCTTCGAGACCGCGCGCTCTCTGATAGCCACACCGCGAG 25405
Qy 1387 erThrAspGlySerMetAsnSerTyrSer----- 1396
Db 25406 AGCTCTCTCCGCGCTCGACTGCTCGCGCCAGGCAAGCGCGCGACCGCTCTCTG 25465
Qy 1397 -----SerGluGlyAsnLeu-----IlePheProGlyVal----- 1406
Db 25466 GACGAGCGGAGGAGCAACGCAAGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25525
Qy 1406 ----- 1406
Db 25526 AAGGATGGCCCTCTCGCTGCTGACTCTCTCGCGCGCTCTCTCTCTCTCTCTCTCTCTCT 25585
Qy 1407 -----ArgLeuAlaSerAspSerGlnPheSer----- 1415
Db 25586 GCGAGCGCGCT 25645
Qy 1416 -----AspPheLeuAsp----- 1419
Db 25646 AGGCGCGCT 25705
Qy 1420 -----GlyLeuGlyProAlaGlnLeuValGlyArg 1430
Db 25706 TCTCTCTGGCGCT 25764
Qy 1430 lnThrLeuAlaThrProAlaMetGlyAspIleGln-----ValGlyMetMetA 1446
Db 25765 -----AGTCAGGCGAGATCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25807
Qy 1446 spLysLysGlyGlnLeuGluValGluIleAlaArgAlaArgGlyLeuValLysProG 1466
Db 25808 CCCTCGAGAGCGCGCGCATCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 25867
Qy 1466 lySerLysThr-----LeuProAlaProTyrValLysValTyrLeuLeuA 1481
Db 25868 GCAACGGGCGCATGGCGCGCT 25927
Qy 1481 spAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeu----- 1498
Db 25928 CCTGGGGCGACAGGCTCTCCATCGCGCGCTCAACAGCCCCCAGCGGCGCTCTCTCTCTCT 25987

QY 1499 --GluPro-----LeuTyrGlnGlnLeuLeuSerPheGluGluSerProG 1513
Db 25988 GCGAGCCCGCGCCATCGAGCGCTGATCGCTCAGCGAGCGAG----- 26037
QY 1513 LnglyArgValLeuGlnIleValTrpGlyAspTyrGlyArgMetAspHisLysSerP 1533
Db 26038 -----GTCTTCGCCCGAAGTCCGCGTCTGACTAGCGCTCC 26074
QY 1533 heMetGlyValAlaGlnIleLeuLeuAspGluLeuGlu-----LeuSerAsnMet-Val 1550
Db 26075 ACTCCGCCAGATGGAGCGCTGACAGAGCTCGCGGAGGTCTAGCCAAATCGCTC 26134
QY 1551 IleGlyTrpPheLysLeuPheProSerSerLeu---ValAspProThrSerAlaPro 1569
Db 26135 CTGCGA-----CGTGGAGCTCCCTCTTTATTGACCGTCAACCGGACCA 26179
QY 1570 LeuThrArgAlaSer-----GlnSerSerLeuGluSerSerThrGlyProSer 1586
Db 26180 GGCTCGAGCGCTCGAGCTCGAGCGCGCTACTGGTATCGAAACCTCCGGCAAAACGCTCC 26239
QY 1587 TyrSerArgSer 1590
Db 26240 TGTTCTCGAGCG 26251

RESULT 15
US-08-729-214-6
; Sequence 6, Application US/08729214
; Patent No. 5817502
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Hammer, Phillip B.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kinner, Sabine
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,214
; FILING DATE: TEA
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1506/CIPS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-729-214-6

Alignment Scores:

Pred. No.: 1,36e-07 Length: 28958
Score: 270.50 Matches: 436
Percent Similarity: 30.17% Conservative: 194
Best Local Similarity: 20.88% Mismatches: 768
Query Match: 3.26% Indels: 696
DB: 1 Gaps: 90
US-09-617-099B-1 (1-1590) x US-08-729-214-6 (1-28958)
QY 3 AlaProLeuGlyProArg-----GlyArgProAlaProThrProAlaAlaSerGlnPro 20
Db 20577 GCTCTGTGCGCCAGAGCGAGCGCGCTGCGCGCCAGCCAGAGCTCCGACCA 20636
QY 21 ProGlnProGluMetProAspLeuSerHisLeuThrGluGluGlnArgLysIleLe 40
Db 20637 CTTCTCTGCGCCACGACGACT-----CGCCCTGCGCGATGTAGCCTA 20678
QY 41 LeuAlaValMetAspArgGlnLysLysGluGluGlnLysGlnSerValLeuLysIle 60
Db 20679 CTGCGT----- 20694
QY 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
Db 20685 ---CGCCACACCCCGCGCTA-CCTTCGAGCACCGTGCCTCTCTGCTCCACGACCGCG 20740
QY 81 ValAsn-----AsnValLeuGlnProGlnGlnLysGlnPro----- 92
Db 20741 AAGAGCTCTCTCCGCGCTCGATTCTGCTCGCCAGGAGCGCCCGCCGCGCGCTCG 20800
QY 93 ---AsnGluLysGluProGlnThrLysLeuHisGlnGlnPheGlnMetTyrLysGluGln 111
Db 20801 TCGACGAGCGAGACCGACCGACGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20842
QY 112 ValLysLysMetGlyGluGlnSerGlnGlnGlnGlnGlnLysGlyAspAlaProThr 131
Db 20843 ---CTGCGC-----AAGGCTCGCAGTGGAGGAGGAGTGGCCCTCTCC 20881
QY 132 CysGlyIleCysHisLysThrLysPheAlaAspGlyCysGlyHisAsnCysSerTyrCys 151
Db 20882 TCTCGATACCTCCGCGGTCTTCC-----GGGCACAGCTCGAAGCGTGG 20926
QY 152 Gln-----ThrLysPheCysAlaAspGlyCly-----Arg 162
Db 20927 AGCGCGCCTCGCGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 20986
QY 163 ValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGlnGlu 182
Db 20987 GCGCGCCCGCTCGACCGGGTCCGAGTGTCCAGCCCGCGCTGTCTCGATGATGCT 21046
QY 183 IleLeuThrLysSerGlyAla----- 189
Db 21047 CGCTGCGCCCGCTGTGCGCGCTCCATGGGCGTCCAGCCCGCGCTGCTGCGCCATAGCC 21106
QY 190 -----TrpPheTyrAsnSerGlySerAsnThrLeu-GlnG 201
Db 21107 AGGCGAGATCGCCCGCGCTGTGTGGCGGGCGCGCTGTGCTCGAGAGCTGCGCAAGC 21166
QY 201 nProAspGlnLysValPro-----ArgGlyLeuArgAsnGluGluAlaProGlnGlu 219
Db 21167 TGGTGGCGTGGCGAGCGCTGCGCTGCTGCGAGCTCGCGCGCGCGCGCGCGCGCG 21215
QY 219 sLysAlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerVal 239
Db 21216 -----CATGGCGCGGTGGAGCTGCGCGAGCGCGAGTCCGACCGCGCGCTCC 21262
QY 239 oAlaValGluLysGlyArgAlaHisGlyLeuThrArgGln-----AspThrI 255
Db 21263 AGCGCTATGGGATCGCGCTCTCCATCGGGGCGGATCAACAGCCCTCGTTTTCAGCAGATC 21322
QY 255 eLysAsnGly-----SerGlyValLysHisGlnIle-- 265
Db 21323 CGGCGAGCCCGCTGCGCGCTGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCGCTGCTG 21382

```
QY 1387 erThrAspGlySerMetAsnSerTyrSer----- 1396
Db 25406 AGTCCTCTCTCGCGCTCGACTCGCCGAGACAAAGCCGCCGAGCACCCTGCTCG 25465
QY 1397 -----SerGluGlyAsnLeu-----IlePheProGlyVal----- 1406
Db 25466 GACGAGCGGAAGACCAAGCAGCTGCTTCTGCTCTTCTGCGCAAGGCTGCGAGTGG 25525
QY 1406 ----- 1406
Db 25526 AAGGGATGGCCCTCTCTGCTGCTGACTCTCGCCCGCTCTTCGCGCTCAGTCGAAGCAT 25585
QY 1407 -----ArgLeuAlaSerAspSerGlnPheSer----- 1415
Db 25586 GCGAGCGCGCTCGCTCTCTCAGCTGAGTGGAGCTGCTGCGCGTCTGCGCGCGAGCG 25645
QY 1416 -----AspPheLeuAsp----- 1419
Db 25646 AGGGCGCCCTCTCTCGACCGCGCTGCTGACAGCCCGCTCTTTGCGCTCATGG 25705
QY 1420 -----GlyLeuGlyProAlaGlnLeuValGlyArgG 1430
Db 25706 TCTCCCTGCGCGCTCTGCGCTCGCTGAGCGCGCTGAGCGCGCTGCGCGCAC- 25764
QY 1430 lnThrLeuAlaThrProAlaMetGlyAspIleGln-----ValGlyMetMetA 1446
Db 25765 -----AGTCAGGCGGAGATCGCGCGCTCTGTCGAGCGGCTCTCT 25807
QY 1446 splyslysGlyGlnLeuGluValGluIleIleArgAlaArgGlyLeuValValysProG 1466
Db 25808 CCTCGAGGACGCGCGCTCGCTGCGCGCTGCGAGCAAGCGCTCACCGCTCGCGCG 25867
QY 1466 lySerlysthr-----LeuProAlaProTyrVallylsValTyrLeuLeuA 1481
Db 25868 GCACGGGCGCATGGCGCGCTCGAGCTCGCGCGCTCCGACCTCCAGACCTACCTCGCTC 25927
QY 1481 spAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeu----- 1498
Db 25928 CTTGGGGGACAGGCTCTCCATCGCGCGCTCAACAGCCCGGCGGCGCTGCTGTCGG 25987
QY 1499 --GluPro-----LeuTyrGlnGlnLeuSerPheGluGluSerProG 1513
Db 25988 GCGAGCCCGCGCATCGACGCGCTGATCGACTCGCTCACCGCGCGCAG----- 26037
QY 1513 lnGlyArgValLeuGlnIleIleValTyrGlyAspTyrGlyArgMetAspHisLysSerP 1533
Db 26038 -----GTCTTCGCGCGGAAAGTCCGCGCTGAGCTAGCGCTCCC 26074
QY 1533 heMetGlyValAlaGlnIleLeuLeuAspGluLeuGlu-----LeuSerAsnMet-Val 1550
Db 26075 ACTCCGCCGATGGAGCGCGTCCAGAGAGCTCGCGCAGGTCTAGCCACATCGCTC 26134
QY 1551 IleGlyTyrPheLysLeuPheProProSerSerLeu---ValAspProThrSerAlaPro 1569
Db 26135 CTCGGA-----CGTGGAGCTCCCTCTTTATTGACCGCTCACCGCGCACCA 26179
QY 1570 LeuThrArgArgAlaSer-----GlnSerSerLeuGluSerSerThrGlyProSer 1586
Db 26180 GGTCCGACGGCTCCGAGCTCGAGCGCGCTACTGGTATCGAAACCTCGGCAACCGCTCC 26239
QY 1587 TyrSerArgSer 1590
Db 26240 TGTTCGAGCG 26251
```

Search completed: November 16, 2003, 15:18:59
Job time : 992 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 06:43:01 ; Search time 850 Seconds
(without alignments)
5049.537 Million cell updates/sec

Title: US-09-617-099B-1
Perfect score: 8285
Sequence: 1 MSAPLGRGRPAFTPAASQP.....TTRASQSSLESSTGSPYSRS 1590

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q=/cgn2_1/USRIO_spo01/US09617099/runat_14112003_191545_23673/app_query.fasta_1.1735
-DB=N_Geneseq 19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09617099 -CGN 1.1.0 @runat_14112003_191545_23673 -NCPUP=6 -ICPUS=3
-NO_WMAP -LARGEQUERY -NGS_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq 19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	8285	100.0	4980	22	AAF86684	Mouse Rim2-encodin
2	2936.5	15.5	6637	22	AA198097	Human neuroblastom
3	1781.5	21.5	1302	19	AAV40485	Human secreted pro
4	1042	12.6	7325	24	AAS94817	Human DNA sequence
5	1015.5	12.3	2276	21	AAC60011	Human secreted pro
6	992	12.0	7395	23	ABL12715	Drosophila melanog
7	598.5	7.2	536	22	ABA11415	Human nervous syst
8	470.5	5.7	20604	23	ABL12714	Drosophila melanog
9	420	5.1	378	21	AAC74934	Human ORFX ORF489
10	374	4.5	9027	22	AAH81795	Human differential
11	368.5	4.4	24387	22	AAO7156	Human reproductive
12	368	4.4	355	20	AAV90372	EST clone DM118
13	368	4.4	355	20	AAV89630	EST clone CO618
14	360.5	4.4	9551	20	AA222301	CDNA encoding a hu
15	311	3.8	6755	19	AAV21511	Staphylococcal bac
16	298	3.6	16857	23	ABL28687	Drosophila melanog
17	298	3.6	18439	23	ABL17911	Drosophila melanog
18	298	3.6	40766	23	ABL28686	Drosophila melanog
19	298	3.6	42429	23	ABL17910	Drosophila melanog
20	296	3.6	9807	23	ABU06949	Drosophila melanog
21	295	3.6	7568	23	AA808823	DNA encoding novel
22	295	3.6	9220	23	ABV24414	Human prostate exp
23	295	3.6	9373	23	ABV25265	Human prostate exp
24	295	3.6	9373	23	ABU07021	Drosophila melanog
25	295	3.6	9416	24	ABK83800	Human cDNA differe
26	295	3.6	12505	23	ABL07020	Drosophila melanog
27	294	3.5	5502	23	ABL28649	Drosophila melanog
28	293	3.5	4509	23	ABL19875	Drosophila melanog
29	293	3.5	4810	23	ABL03815	Drosophila melanog
30	292.5	3.5	7796	23	ABL20262	Drosophila melanog
31	292.5	3.5	9880	21	AZA46851	Nucleotide sequenc
32	292.5	3.5	23914	23	ABL20258	Drosophila melanog
33	292.5	3.5	109519	22	AA508693	Microzoosporea DNA
34	291	3.5	13095	25	ABX07542	S. pneumoniae type
35	291	3.5	13152	25	ABX07541	S. pneumoniae type
36	291	3.5	14328	25	ABX07543	S. pneumoniae type
37	291	3.5	2162598	25	AB556454	Streptococcus pneu
38	290.5	3.5	5796	23	ABL20263	Drosophila melanog
39	290	3.5	4483	19	AAV52496	Streptococcus pneu
40	290	3.5	48300	22	AAF61281	N. magadaii bacter
41	288.5	3.5	33529	23	AA317367	DNA sequence of S.
42	285.5	3.4	6307	23	AA574131	DNA encoding novel
43	284.5	3.4	3819	24	AB211786	Human polynucleoti
44	284.5	3.4	4215	25	ACC50305	Breast cancer asso
45	281.5	3.4	18506	23	ABL05592	Drosophila melanog

ALIGNMENTS

RESULT 1

AAF86684
ID AAF86684 standard; cDNA; 4980 BP.
XX
XX AAF86684;
XX AC
XX
DT 25-JUL-2001 (first entry)
XX
XX Mouse Rim2-encoding cDNA.

XX Mouse; murine; Rim2; Rim isoform; Rab3-interacting molecule;
XX Rab3-dependent intracellular vesicle fusion regulator;
XX GTP/GDP exchange factor; CAMP-GEFII; neuron; endocrine tissue;
XX exocytosis; hormone secretion; neurotransmitter release;
XX presynaptic membrane; neuropathy; central nervous system disorder;
XX brain disorder; secretion disorder; pituitary gland; hypothalamus;
XX pancreatic beta-cell; diabetes; parotid gland; zinc finger; PDZ domain;

KW C2 domain; ss.
 XX Mus musculus.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 32..4804
 FT /*tag= a
 FT /product= "Mouse Rim 2"
 FT /note= "Rim 2 is a novel isoform of Rim"
 XX
 PN EF1090986-A1.
 XX
 PD 11-APR-2001.
 XX
 PF 31-JUL-2000; 2000EP-0116148.
 XX
 PR 08-OCT-1999; 99JP-0288372.
 XX
 PA (JCRP-) JCR PHARM CO LTD.
 PA (SEIN/) SEINO S.
 XX
 PI Seino S, Shibasaki T, Ozaki N;
 XX
 DR WPI; 2001-275168/29.
 DR P-PSDB; AAB73488.
 XX
 XX New Rim 2 proteins and nucleic acids encoding the protein useful in
 PT diagnosing endocrine-related diseases or neuropathy and in developing
 PT agents for the prevention or treatment of such diseases
 XX
 PS Claim 4; Page 19-31; 42pp; English.
 XX
 CC This sequence represents cDNA encoding mouse Rim2. Rim2 is a novel
 CC isoform of Rim (Rab3-interacting molecule, also referred to in the
 CC specification as Rim1), and both proteins contain a zinc finger,
 CC PDZ and two C2 domains. Rim2 is able to specifically interact with
 CC the GTP/GDP exchange factor cAMP-GEFII, a cAMP sensor molecule, and
 CC is believed to be a regulator of Rab3-dependent intracellular vesicle
 CC fusion in neurons and endocrine tissue. It is therefore likely to
 CC play a key role in neurotransmitter release at presynaptic membranes,
 CC and in hormonal secretion. The Rim2 protein is useful in diagnosing
 CC endocrine-related diseases or neuropathy and in developing agents for
 CC the prevention or treatment of such diseases. Rim2-specific
 CC hybridisation probes or antibodies are useful in the diagnosis of brain
 CC or central nervous system disorders, or for diagnosis of secretion
 CC disorders of the pituitary gland, hypothalamus, pancreatic beta-cells
 CC and parotid gland.
 XX
 SQ Sequence 4980 BP; 1503 A; 1161 C; 1282 G; 1034 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0 Length: 4980
 Score: 8285.00 Matches: 1590
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-617-099B-1 (1-1590) x AAF86684 (1-4980)
 QY 1 MetSerAlaProLeuGluProArgGlyArgProAlaProThrProAlaSerGlnPro 20
 DB 32 RTGTGGCTCCGCTCGGGCCCGGGCCCGGGCTCCACCCCGGGCTCTCAACCT 91
 QY 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysIleIle 40
 DB 92 CCTCCGACCCGAGATGCGGACCTCAGCCACCTCAGGAGAGAGAGAGAAATCATC 151
 QY 41 LeuAlaValMetAspArgGlnLysLysGluGluGluGlnSerValLeuLysIle 60
 DB 152 CTGGCTGTCTATGATCGTCAGAGAAAGAGAGAGAGAGAGAGAGAGATC 211
 QY 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80

DB 212 AAAGAGAGACACAAAGCACACACACAGTGGTTTCCCTTAGTGGATCACTGACTG 271
 QY 81 ValAenAenValLeuGlnProGlnGlnLysGlnProAenGluLysGluProGlnThrLys 100
 DB 272 GTAATAACGTTCTGACGCCCCAGCAAAACCAACCAATAGAGAGAGAGAGAGAGAG 331
 QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120
 DB 332 CTGCACCAACCAATTTGAATGTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCGCAG 391
 QY 121 GlnGlnGlnGluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
 DB 392 CAGCAGCAAGATTT 451
 QY 141 AlaAspGlyCysGlyHisAenCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
 DB 452 GCAGATGATGATCGGCCATAATTTGTTCTATTGCCAAACCAAGTTCTGTGCTCGATGTGA 511
 QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAenLeuCysArgLysGln 180
 DB 512 GGTGAGTGTCTTTACGCTCAACCAAGGTATATGTGGTGTGTAAATTTGTGCCGAAACAA 571
 QY 181 GlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAenSerGlySerAenThrLeuGln 200
 DB 572 CAAGAAATCTCTCACTAAATCAGGACATGGTTTTATAATAGTGGTCTTAACACACTGCAG 631
 QY 201 GlnProAspGlnLysValProArgGlyLeuArgAenGluAlaProGlnGlnLysLys 220
 DB 632 CAACCTGATCAAAAGGTTCTCTGAGGGCTTCCAAATGAGGAGAGAGAGAGAGAGAG 691
 QY 221 AlaLysLeuHisGluGlnProGlnPheGlnIleAlaProGlyAspLeuSerValProAla 240
 DB 692 GCAAACATACACAG 751
 QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAenGlySerGly 260
 DB 752 GTTGAGAAAGCCGAGCTCTGGGCTCAACAGAGAGATCTATTAAATGGATCAGAG 811
 QY 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
 DB 812 GTGAAGCACCATGTTGCCAGTGACATGCTTCAGACAGAAACGAAGTCCATCAGTGCC 871
 QY 281 ArgAspGlnAenArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrVal 300
 DB 872 AGGGATCAAAATCGAAGATACGAGCAAAAGTGAAGAGAGAGAGAGAGAGAGATGTT 931
 QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSerGlnArg 320
 DB 932 CCTTCAGATGTGTACAAATGCCAGATCTCTTCGGATTATGCTGATAGAGATCTCAGCGT 991
 QY 321 GluProGlnPheTyrGluGluProGlyHisLeuAenTyrArgAspSerAsnArgArgGly 340
 DB 992 GAGCCTCAATTTATGAAGAACCTGTGTCTATTAATTAACAGGATTTCTTAACAGGAGAG 1051
 QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
 DB 1052 CATAGACATTCCAAAGAGTATATTGTGGATGATGATGATGATGATGATGATGATGAT 1111
 QY 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAenLeuAla 380
 DB 1112 GAAGACAAAGAGAGAGAGAGAGAGATACCAGGCACCTACAGAGAGATGATCCAAATCTGCC 1171
 QY 381 ArgTyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValSer 400
 DB 1172 CGGTATCCGTAAAGCCACACCTACAGAGACAAATGCCATGCCAGCTCAGGTGTGCC 1231
 QY 401 ArgAlaArgHisGluArgHisSerAspValSerLeuAlaAenAlaGluLeuGluAsp 420
 DB 1232 AGGGCAGCAGATGATGAT 1291
 QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440

1292 TCCAGGATTCTCTGCTAAGGATGGATAGACCATCAAGGCAGAAAGATCTGTATCTGAACGT 1351
441 ArgAlaalaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
1352 AGAGCTGCATGGAAACCAACCATGTTATTCATGGAAGAACTCGAGAGGCTCAGGGA 1411
461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
1412 CAAAGTTCTTATCCACAAAGGACCTCAATCATAGTCTCTCCACCCCTCGCGGAGCCCT 1471
481 IleProLeuAspArgProAspMetArgAlaAspSerLeuArgGlyGlnHisLeu 500
1472 ATACCCCTTGATAGACAGACATAGCGCGCTGACTCTCTACGGAACAGCACCACTTA 1531
501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
1532 GATCCAGCTCTCTCTGTAGGAAACAGACGAGAAATAATGGAACCATGTTAAGGAAT 1591
521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
1592 GATTCCTTGAGTTTCAGACAGTCCGAGTCAGTGAGCGCCGCCCCACCAAGGCTCATAA 1651
541 SerLysGlyGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAla 560
1652 TCCAGAAAGAGGAGTAAATCGCCAGGTTTCACTGAGCAGCTCGGAGGAGGAGCTGGCA 1711
561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
1712 TCCACACCTTGAGTATCAAGCTGTGATGATGTGGAGCTGGAACCGAGAGTGTGAGTGAG 1771
581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
1772 AAAGGGGACAGTCARAAAGGGAAGAAAGAACTAGTGAGCAGGAGTTTGTTCGGATTCT 1831
601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
1832 AACACCAAGCTCTGAGAGACAAAGAAAGAGATGTAATGTTGGTGGCCACTCTTTGGAAGAG 1891
621 AspLeuGluTrpSerGluProGlnLysLysAspSerGlyValAspThrCysSerSerThr 640
1892 GATTTGGATGGTCTCAGAGCTCAGATTAAAGGACTCTGGGGTAGATACCTGTAGTAGACA 1951
641 ThrLeuAsnGluLysHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
1952 ACCCTTAAAGAGAGCATAGCCATAGTAGTAGACACCTGTGACCTGGCAGCCATCCAA 2011
661 AspGlyAspArgLeuIleGlyArgGlyLeuLeuAsnLysArgLeuLysAspGlySerVal 680
2012 GATGGAGATCGCCTAATTGGTCTGATTTTATTAAATAAGCGTTTAAAGATGGAGTGTA 2071
681 ProArgAspSerGlyAlaMetLeuGlyLysValValGlyGlyLysMetThrGluSer 700
2072 CTTGAGACTCAGGAGCAATGCTGGGCTTAAAGGTTGTAGGAGGAAAGATGACTGAATCA 2131
701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
2132 GGTGCACTTTGTGCAATTTATCCAAAGTAAAGGAAAGGAGTTTACTGATACCTGAGGA 2191
721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr 740
2192 CATCTTAGACCAAGTGTAGTCTTGGATGGAATGGAGGCTATTGCAAGAGGCCACA 2251
741 PheGluGluValTyrAsnIleLeuLeuSerLysProGluProGlnValGluLeuVal 760
2252 TTTGAGGAAGTTTACAACTATTCTAGAAATCCAAACCTGAAACCAAGTTGAGCTTGT 2311
761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
2312 GTTTCAGGCCAATTGGAGATATTCCTAGAATACTCTGATAGCAGCATGCACAACTGGAA 2371
781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
2372 TCCAGTCTTAGCTCATTTGAATCTCAAAAATAAGGACCGCTCTTCTATATCCGTTACTCA 2431

801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
2432 CCCATGAGTCTTGCGATGCTGAGGATGTCGCCGAGTTCTTATCTGACAGCTTTCATA 2491
821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
2492 AAATATGTTTGACAAAGTTGTCACCAAGTTGATAGTTACAAATTTTGGAGCAAGGAT 2551
841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
2552 CTCCCTTCCAGGCAAGATGGGAGGCCAGGAATCCTTATGTTAAGATTTACTTCTTCCA 2611
861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
2612 GATAGAGTGATAAATAAGAGAGAAACAAACAGTCAGAAACCTTTGGAAACCCAA 2671
881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
2672 TGGAACCAAGACTTTCATTTATCTCTGTCCACCGAAGAGAAATTCGTGACAGATGCTG 2731
901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGlu 920
2732 GAAATTCACCTTTGGGATCAAGCTAGAGTTCGAGAAGAGAGAGCGAATTCCTAGAGAG 2791
921 IleLeuIleGluLeuGluThrAlaLeuAspAspGluProHisTyrTyrLysLeuGln 940
2792 ATTTAATTGAATTGAAACAGCTTTGCTAGATGATGAGCCGACACTGGTATTAAGCTGAG 2851
941 ThrHisAspValSerSerLeuProLeuProAspProSerProTyrLeuProArgArgGln 960
2852 ACCCATGATGTCCTCTCATTTGTCATCTCCCTCGCCCTTCCCCATATCTGCCCGAGGAG 2911
961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
2912 CTCCTATGGAGAGAGCCCAACGCGCAGGCTGCAAGTGCGAAGAGAAATTAAGTCACAGTGA 2971
981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisLeuGly 1000
2972 GTGCTCTGACTACGACTGCGAGGATGCGTGGGAGTAGTGTCCAGATTTATGCACAAATGC 3031
1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
3032 CGCGATCTTCAAAAGTCTCCAGTTCGCTGCCAGCAACAAAGTCATGTCAATCAATCTGC 3091
1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
3092 TCACCATCAGGCTCTCTCTCATCGAGTAGATGTTATAGGAAGCAAGGTCATGTCGCTCT 3151
1041 SerAlaProProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
3152 AGTGGCCCTCTCTCTCAAGGAATGGAACAGAGGCGCACCGAGGAGCACGTCCTACTGCG 3211
1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
3212 CATTAACACAAATTAGCCGAATGATAGACACCGTGTATGATGACCACTACTCTTCTCA 3271
1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
3272 GATAGAGCAGGATTTGGAACGACAGATAGACAGCCATATCAAGATCCAGATCAACA 3331
1101 GluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProAsp 1120
3332 GAACCAACGCTCTCTCTAGAGCGGACCAACCCGCTCCAGATCTCTGAACTCTCTGAT 3391
1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140
3392 ACAACACCTCATGAGTGTGCTTCAATTAAGTGAAGATCTGCCCTCTCTTCACT 3451
1141 AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrPro 1160
3452 GCCTTATCGAGGTCTCACCTCTGTAACCGGTCTGTCCAGAACGCCCATCAAGTACTCG 3511

Qy 735 LeuLeuGlnGlyAlaThrPheGluGluValTrpAsnIleLeuLeuSerLysProGlu 754
Db 2274 CCCCTGCCGGGAGCTACAAATGAAGAAGTTTACACATTTATTTAGAAATCAAAATCAGAA 2333
Qy 755 ProGlnValGluLeuValSerArgProIleGlyAspIleProArgLysProAspSer 774
Db 2334 CCTCAAGTTGAATTTATTTGTTTCAAGGCTTATGTTGACATCCCGGATTCCTGAGAGC 2393
Qy 775 ThrHisAlaGlnLeuLeuSerSerSerPheGluSerGlnLysMetAspArgPro 794
Db 2394 TCCCAACCTCCACTGGAGTCAGTTCAAGTTCTTTGAATCTCAGAAGATGGAAGGCT 2453
Qy 795 SerLeuSerValThrSerProMetSerProGlyMetLeuArgAspValProGlnPheIleu 814
Db 2454 TCCATTTCTGTATTCTTCCAAACAGTCTCGAGCTCTAAAGATGCCCCCAAGTCTTA 2513
Qy 815 SerGlyGlnLeuSerIleLeuLeuTrpPheAspLysValGlyHisGlnLeuIleValThr 834
Db 2514 CCAGGGCAACTTTCTGTGAAGTTGTGGTATGATTAAGTGGACACACAGCTGATTTGTAAT 2573
Qy 835 IleLeuGlyAlaLysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrVal 854
Db 2574 GTTCTGCAAGCAACAGATCTACCTGCTAGATAGATGACGCTCTCGAAATCCCTATGTA 2633
Qy 855 LysIleTyrPheLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValLys 874
Db 2634 AAAATGATTTCTTCCAGATAGAGTGTATAAAGTAAAGGAGGAGCAACAAACAGTAAG 2693
Qy 875 LysThrLeuGluProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGlu 894
Db 2694 AAATACATAGAACCAAAATGGAATCAACTTTGTCTATTTCATGTACATCTGTAGAT 2753
Qy 895 PheArgGluArgMetLeuGluIleThrLeuTrpAspGlnAlaArgValArgGluGluGlu 914
Db 2754 TTTAGAGAACCAATGTTAGAAATACTGTGTGGACCAACCAAGAGTCAAGAAGAGAA 2813
Qy 915 SerGluPheLeuGlyGluIleLeuIleGluLeuThrAlaLeuLeuAspAspGluPro 934
Db 2814 AGTGAATTTCTTGGAGAGATCTCATAGAAATGGAGACAGCGCTTTTAGATGATGAACCG 2873
Qy 935 HisTrpTyrLysLeuGlnThrHisAspValSerLeuProLeuProArgProSerPro 954
Db 2874 CATTTGGTATAAATCTCAGACATGATGATGCTTCTCATTACCTCTCGCTCAGCAATCCT 2933
Qy 955 TyrLeuProArgArgGlnLeuHisGlyGluSerProThrArgArgLeuGlnArgSerLys 974
Db 2934 TTCATGCCCAAGCGACATATTTCTAGGAGAAAGCTTAGCAAAAGCTACAAAGATCTCAG 2993
Qy 975 ArgIleSerAspSerGluValSerAspTyrAspCysGluAspGlyValGlyValValSer 994
Db 2994 CGAATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3053
Qy 995 -----AspTyrArgHisAsnGlyArgAspLeuGlnSerSerThrLeuSerValProGlu 1012
Db 3054 CAGTAGGCTATAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCT 3113
Qy 1013 GlnValMetSerSerAsnHisCysSerProSerGlySerProHisArgValAspValIle 1032
Db 3114 CAGCAAGACCAACTCATCCGCTCAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAG 3173
Qy 1033 GlyArgThrArgSerTrpSerProSerAlaProProGlnArgAsnValGluGlnGly 1052
Db 3174 GGAAGCCGCTTACGTTTACCAATGTG---CCATTACAGAGGAGTTTAGTGAATTA 3230
Qy 1053 His-----ArgGlyThrArgAlaThrGlyHisTyrAsnThrIleSerArgMet----- 1068
Db 3231 CATCCAAACAGAGGTCGCTTCCAAACCAACAGATGATGCTCTCCGAGATCGAGTT 3290
Qy 1069 ---AspArgHisArgValMetAspAspHisTyrSerSerAspArgAspArgCysGlu 1087
Db 3291 GATCATAGAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3350

Qy 1088 AlaAlaAspArgGlnProTyrHisArgSerArgSerThrGlu 1101
Db 3351 ATGCTGCCAGACA-----AAACGAGGACGAATGTCAGAA 3386
RESULT 3
AAV40485
ID AAV40485 standard; cDNA; 1302 BP.
XX AAV40485;
AC AAV40485;
XX 09-NOV-1998 (first entry)
XX Human secreted protein CO618_1 cDNA.
XX Secreted protein; CO618_1; human; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 123..1247
XX /*tag= a
XX MO9831802-A1.
XX 23-JUL-1998.
XX 21-JAN-1998; 98WO-US01007.
XX 20-JAN-1998; 98US-0010047.
XX 16-JUN-1997; 97US-0072103.
XX 16-JUN-1997; 97US-0877035.
XX (GENY) GENETICS INST INC.
XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
XX Racie LA, Spaulding V, Treacy M;
XX WPI; 1998-414101/35.
XX P-FSDB; AAW29640.
XX New isolated polynucleotides and secreted proteins - obtained from
XX human foetal kidney, adult testes, placenta, adult brain and foetal
XX brain cDNA libraries
XX Claim 22; Page 71-72; 104pp; English.
XX This full-length cDNA clone, designated CO618_1 and deposited at
XX ATCC 98296, codes for a novel human secreted protein (see AAW29640).
XX It was isolated from a human adult brain cDNA library using
XX methods which are selective for cDNAs encoding secreted proteins,
XX and was identified as encoding a secreted or transmembrane protein
XX on the basis of computer analysis of the amino acid sequence of the
XX encoded protein. CO618_1 shows some similarity to human male bone
XX marrow myeloblast mRNA for KIAA0237 protein. Novel secreted
XX protein-encoding polynucleotides (see AAV40479-90) have been
XX identified in foetal kidney, adult testis, placenta, adult brain
XX and foetal brain cDNA libraries. These can be used in the
XX production of recombinant polypeptides in host (preferably
XX mammalian) cells. The polypeptides may have e.g. nutritional
XX activity, cytokine and cell proliferation/differentiation activity,
XX immunostimulant or immunosuppressive activity, haematopoiesis
XX regulating activity, tissue growth activity, activin/inhibin
XX activity, chemotactic or chemokinetic activity, haemostatic or
XX thrombolytic activity, receptor/ligand activity, antiinflammatory
XX activity, cadherin or tumour invasion suppressor activity, tumour
XX inhibition activity, or other activities.
XX Sequence 1302 BP; 404 A; 286 C; 322 G; 286 T; 4 other;
SQ
Alignment Scores:
Pred. No.: 3,5e-78 Length: 1302
Score: 1781.50 Matches: 359
Percent Similarity: 88.41% Conservative: 7

Best Local Similarity: 86.71% Mismatches: 7
 Query Match: 21.50% Indels: 41
 DB: 19 Gaps: 1

US-09-617-099B-1 (1-1590) x AAV40485 (1-1302)

QY 1218 GlyTyrAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerVal 1237
 DB 3 GGATGGGATCCTCATAGAGGGCGAGTATGTTTCTACTAATCTTCGACAGTGTATGTA 62
 QY 1238 SerAspValSerAlaValSerArgThrSerAlaSerArgPheSerThrSerTyr 1257
 DB 63 AGTGATATATCTCGGGTTTCAAGGACTAGTAGTCTCTCGTTTCAGCAGCACAAGCTAC 122
 QY 1258 MetSerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLys 1277
 DB 123 ATGTCTGTCTCAATCAAGACGCCAGGAGGAACAAGAAATCAGTGTCTTTTATCATCCAAA 182
 QY 1278 MetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSer 1297
 DB 183 ATGCAAGCAGACAAATGGGCATATCAGGAGAGACATGACAAAAGACACGATCATCT 242
 QY 1298 GlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAla 1317
 DB 243 CGAGACATGTGCTCACTGAGAGAAATGATGGCAGCAGTCTGACACACTGCAGTGGGCACC 302
 QY 1318 LeuGlyThrSerGlyLysArgArgSerSerIleGlyAlaLysMetValAlaIleVal 1337
 DB 303 TTGGGCACAGTGGCAAAAGCGCGCTCTAGCCTTGGTGGCCAAAATGTTAGTATCGT 362
 QY 1338 GlyLeuSerArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLys 1357
 DB 363 GGTCTGTCTACGGAAGTGCAGTGTCTTCTAGCTCAGCCAAACGGAACGAGAGTTAA 422
 QY 1358 LysLeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsn 1377
 DB 423 AAATTAAGCAGCACTCTCCAAAGAGTACAGAAACAGGCGCTGGCCCTGGAAATGAGGAAC 482
 QY 1378 TrpMetThrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSer 1397
 DB 483 TGGATGACTCGACAGCAAGCCGAGAGTCTACAGATGGTAGTGAACACGCTACAGCTCA 542
 QY 1398 GluGlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPhe 1417
 DB 543 GAAGGAATCTGATTTCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
 QY 1418 LeuAspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMet 1437
 DB 603 CTGGATGGCTTGGCCCTGCTCAGTGTGGAGCGCCAGACTCTGGCAACACCTGCAATG 662
 QY 1438 GlyAspIleGlnValGlyMetMetAspLysLysGlyGlnLeuValGluIleIleArg 1457
 DB 663 GGTGATTCAGTAGAGATGATGGAACAAAAGGAGCAGCTGGAGGTAGAAATCATCCGG 722
 QY 1458 AlaArgGlyLeuValLysProGlySerLysThrLeuProAlaProTyrValLysVal 1477
 DB 723 GCCGTGGCTGTTGTTTAAACACAGGTTCCAGACACTGCCAGCACCTGATGTAAAGTG 782
 QY 1478 TyrLeuLeuAsnGlyValCys----- 1485
 DB 783 TATCTATTAGATAACGGAGTCTGCTAGCCAAAAGGAGCAGCTGGAGGTAGAAATCATC 842
 QY 1485 ----- 1485
 DB 843 CGGGCCCGTGGCCTTGTGTAAACACAGGTTCCAGACACTGCCAGCACCGGTATGTAAAA 902
 QY 1486 ----- 1486
 DB 903 GTGTATCTATTAGATAACGGAGTCTGSAATAGCCAAAAGAAACAAAGTGGCAAGAAA 962
 QY 1497 ThrLeuGluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgVal 1516
 DB 963 ACCTGGACCCCTTTTACGACGAGTATATATCTTTTTCGAGAGAGTCTCAAGGAAAGTT 1022

QY 1517 LeuGlnIleIleValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyVal 1536
 DB 1023 TTACAGATCATCGTCTGGGAGATTATGCCCATGATCATCAATCTTTATGGAGTG 1082
 QY 1537 AlaGlnIleLeuLeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeu 1556
 DB 1083 GCCCAGATACCTTTTATAGTGAACCTAGAGCTATCCATATATGGTATCGATGGTTCAACTT 1142
 QY 1557 PheProProSerSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGln 1576
 DB 1143 TTCCACACTTCTCTCCTAGTATCAACCTTGGCCCTCTGACACGAGAGCTTCCCAA 1202
 QY 1577 SerSerLeuGluSerSerThrGlyProSerTyrSerArgSer 1590
 DB 1203 TCATCTCTGAAAGTTCAACTGGACCTTCTTACTTTCGTTC 1244

RESULT 4

ID AAS94817
 AAS94817 standard; DNA; 7325 BP.

XX AAS94817;

DT 14-FEB-2002 (first entry)

DE Human DNA sequence #72 expressed during foam cell differentiation.

XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;
 KW cardiovascular disorder; coronary artery disease; gene therapy; ds.
 XX Homo sapiens.

XX WO200177389-A2.

XX 18-OCT-2001.

XX 04-APR-2001; 2001WO-US11128.

XX 05-APR-2000; 2000US-195106P.

XX (INCY-) INCYTE GENOMICS INC.

XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
 PI Tai J;

XX WPI; 2002-010925/01.

XX Composition useful for diagnosis of conditions, disorders or diseases
 PT associated with atherosclerosis, comprises several polynucleotides that
 PT are differentially expressed in foam cell development -
 XX Claim 1; Page 123-125; 315pp; English.

XX The present invention relates to the isolation of human polynucleotide
 CC sequences that are differentially expressed during foam cell
 CC differentiation. The polynucleotide sequences of the invention or a
 CC composition comprising these polynucleotides are useful as a high
 CC throughput method for detecting altered expression of one or more
 CC polynucleotides in a sample. The polynucleotides can be used in the
 CC diagnosis of disorders associated with foam cell development such as
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
 CC coronary artery disease. The polynucleotide sequences can also be used
 CC as PCR primers and probes. The polynucleotides of the invention are also
 CC useful in gene therapy. AAS94746-AAS95021 represent the human
 CC polynucleotide sequences of the invention which are differentially
 CC expressed during foam cell differentiation.

SQ Sequence 7325 BP; 1526 A; 1985 C; 2033 G; 1778 T; 3 other;

Alignment Scores:

Pred. No.: 1,95e-41 Length: 7325
 Score: 1042.00 Matches: 245
 Percent Similarity: 62.65% Conservative: 62

Best Local Similarity:	50.00%	Mismatches:	113
Query Match:	12.58%	Indels:	70
DB:	24	Gaps:	13
US-09-617-099B-1 (1-1590) x AAS94817 (1-7325)			
Qy	1125	ArgSerMetPro-----	SerLeuMetThrGlyArgSerAlaProProSerProAla 1141
Db	66	CGGCGGATGCCCGCGTGGCTCCCTC-----	GCCGCCCTCCCGCGCGCG 113
Qy	1142	--LeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerThrPro 1160	
Db	114	TGCCTATCCACTCGAGTCCGCCGACCTGGG-----	GCCGGCGCGCGCTACTGCG 167
Qy	1161	GlyThr-----	GlyArgGlyArgGlnLeuProGlnLeu-----
Db	168	GTTTCGGCGGGCGGGTCCCGGGGCGACACCTGCGCGCTTGGAGCGCGCTCGG---	224
Qy	1177	ThrLeuGluArgSerAlaMetAspIleGluArgAsnArgGlnMetLysLeuAsnLys 1196	
Db	225	CCTGTGGAGGCCCTCCCTG-----	245
Qy	1197	TyrLysGlnValAlaGlySerAspProArgLeuGlnAspTyrHisSerLysTyrArg 1216	
Db	246	-----	TCTGGACCCCGCGCCACCTCCGACCCCTTTATCATCATGCG 287
Qy	1217	SerGlyTyrAspProHisArgGlyAlaAspThrValSerThrLysSer-----	Ser 1233
Db	288	CTCCTCTGGAGCTGCGCTGATTCCTTCCCTCATTTCTTGAATTTGGTTTGGC 347	
Qy	1234	AspSerAspVal-SerAspValSer-----	AlaValSerAspThrSerSe 1248
Db	348	AGAAGTCAATTGAAGCCCTTGTGCAAAATCCCTAGGGGTGTCTGTGGAGGCGACCC 407	
Qy	1248	rAlaSerArgPheSerThrSerTyrMetSerValGlnSerGluArgPro-----	1265
Db	408	TGTGATGGGACACACCGAGCTCAGATTTCATGATTCGAGCTGCCTGAGCGCTGCCACCC 467	
Qy	1266	-----	ArgGlyAsnArgLysIleSerValPhe-ThrSerLysMetGlnA 1280
Db	468	CGGGGACCATGTTTAAACGGGAGCCAGGTCTCTGCTCATCTGGGGCTCC-----	517
Qy	1280	snArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspM 1300	
Db	518	-----	AGGAATGTGTGCGGAGCTCCAGCATTCGCGTGAA 554
Qy	1300	etCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyT 1320	
Db	555	TCTGCGATCCAGACCGCGGGG-----	GGGGCTGGGACCA 593
Qy	1320	hrSerGlyLysArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuS 1340	
Db	594	CCACCGCAAGAGCGCGGAGCAGCTGGGTGCCAAGATGTGGCCATCGTGGCGCTGA 653	
Qy	1340	erArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyLysIleLeuA 1360	
Db	654	CTCAGTGGAGCAAGAGCACATCCAGCTTCGACGCTGAAGGGGCCCAAGAGCTGC 713	
Qy	1360	rgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetT 1380	
Db	714	GCAGCAACATCCCGCGGAGCACGAGACAGGATCGCGGTGGAGATGCGGACCGGTC 773	
Qy	1380	hrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyA 1400	
Db	774	CACGCGAGGCGAGCGGGAGTCCACCGATGGGAGCACCAACAGACGCTCCGACGCA 833	
Qy	1400	snLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspG 1420	
Db	834	CGTTTCATCTTCCCACTACCGGCTAGGGCTGAAAGCCAGTTCAGCGATTTCTCGATG 893	
Qy	1420	lyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspI 1440	
Db	894	GGCTGGGACCACTCAGATTGTGGGGCGACACACTGGCAACACCCATGGGAGATG 953	
Qy	1440	leGlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleLeuArgAlaArg 1460	
Db	954	TGCACATTGCCATCGACCGGAGTGGCGAGCTGGAGGTGAAGTATTGAAGCTCGGG 1013	
Qy	1460	lyLeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuL 1480	
Db	1014	GCTTGACCCCAACAGGCTCCAAATCCCTCCAGCCACCTATATCAAGTTTACTGCG 1073	
Qy	1480	euAspAsnGlyValCysIleAlaLysLysThrLysValAlaLysLysThrLeuGluP 1500	
Db	1074	TGAGAAATGGGCGCTGCTTGGCCACAGAGAGACAAAGATGACCAAGAGACCTGTGATC 1133	
Qy	1500	roLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGluIleI 1520	
Db	1134	CCCTGTACACAGCAGGCTCTCTTTTACAGAGGACCCAGGCAAGGTCTCGAGGTGA 1193	
Qy	1520	leValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleL 1540	
Db	1194	TGCTCTGGGAGACTATGGCGCATGGACACCAAGTCTTCATGGGCATGGCCAGATCA 1253	
Qy	1540	euLeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPheProS 1560	
Db	1254	TGCTGGAGGAGCTGGACCTCAGCGCGCGGTCCACCGCTGGTACAAACTCTTCCACCT 1313	
Qy	1560	erSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuG 1580	
Db	1314	CCTCAGTGGCAGACTCCACACTCGGATCCCTCAGGCGCCTGTCCCGTCTTCCCTGG 1373	
Qy	1580	luSerSerThrGlyProSerTyr-Ser 1588	
Db	1374	AGAGTGGCACCGCCCTCATGTCT 1399	
RESULT 5			
ID	AAC60011	standard; cDNA; 2276 BP.	
XX	AAC60011;		
AC	AAC60011;		
DT	26-JAN-2001	(first entry)	
XX	Human secreted protein gene 46 SEQ ID NO:56.		
DB	Human; secreted protein; diagnosis; cytostatic; immunosuppressive; n		
KW	notropic; neuroprotective; antiviral; antiallergic; hepatotropic; a		
KW	ntidiabetic; antiinflammatory; antitumor; antitumor; antitumor; antitumor;		
KW	antibacterial; antifungal; antiparasitic; cardiatic; gene therapy;		
KW	cancer; immune disorder; cardiovascular disorder; wound healing;		
KW	neurological disease; infectious disease; chromosome identification;		
KW	Chromosome 6; ss.		
XX	Homo sapiens.		
OS	WO200058356-A1.		
XX	05-OCT-2000.		
PD	22-MAR-2000; 2000WO-US07535.		
XX	26-MAR-1999; 99US-0126511.		
PR	17-DEC-1999; 99US-0172413.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
PA	Rosen CA, Ruben SM, Komatsoulis G;		
PI	WPI; 2000-594639/56.		
DR	P-PSDB; AAB34818.		
DR	Fifty nucleic acid molecules encoding human secreted proteins, useful		
PT	in the prevention, treatment and diagnosis of cancer, immune disorders,		
PT	cardiovascular disorders and neurological diseases -		
XX			

Claim 1; Page 372-373; 425pp; English.

The polynucleotide sequences given in AAC59966 to AAC60015 encode the human secreted proteins given in AAB34773 to AAB34822. AAB34823 to AAB34952 represent human secreted polypeptide sequences and proteins homologous to them, which are given in the exemplification of the present invention. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral; anti-allergic; hepatotropic; antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; and cardiant. The polynucleotides and polypeptides are useful for preventing, treating or ameliorating a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. The polynucleotides are useful for chromosome identification. They are also useful as probes for diagnosing a disorder related to the female reproductive system, particularly breast and/or ovary cancer. They are also useful in the therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies, agonists and antagonists from the present invention are useful in the diagnosis, treatment and prevention of cancer, immune disorders, cardiovascular disorders, wound healing, neurological diseases and infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence used in the exemplification of the present invention.

Sequence 2276 BP: 699 A: 451 C: 436 G: 689 T: 1 other:

Alignment Scores:

Pred. No.:	1.1e-40	Length:	2276
Score:	1015.50	Matches:	197
Percent Similarity:	90.76%	Conservative:	19
Best Local Similarity:	82.77%	Mismatches:	21
Query Match:	12.26%	Indels:	1
DB:	21	Gaps:	1

US-09-617-099B-1 (1-1590) x AAC60011 (1-2276)

1353	QY	GlulGlyGlyGlyLysLysLeuA	gSerThrValGlnArgSerThrGluThrGlyLeuAla	1372																																																																																																																																																																																																																																																																																																																																																																																																																																																										
		1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494	1495	1496	1497	1498	1499	1500	1501	1502	1503	1504	1505	1506	1507	1508	1509	1510	1511	1512	1513	1514	1515	1516	1517	1518	1519	1520	1521	1522	1523	1524	1525	1526	1527	1528	1529	1530	1531	1532	1533	1534	1535	1536	1537	1538	1539	1540	1541	1542	1543	1544	1545	1546	1547	1548	1549	1550	1551	1552	1553	1554	1555	1556	1557	1558	1559	1560	1561	1562	1563	1564	1565	1566	1567	1568	1569	1570	1571	1572	1573	1574	1575	1576	1577	1578	1579	1580	1581	1582	1583	1584	1585	1586	1587	1588	1589	1590	1591	1592	1593	1594	1595	1596	1597	1598	1599	1600	1601	1602	1603	1604	1605	1606	1607	1608	1609	1610	1611	1612	1613	1614	1615	1616	1617	1618	1619	1620	1621	1622	1623	1624	1625	1626	1627	1628	1629	1630	1631	1632	1633	1634	1635	1636	1637	1638	1639	1640	1641	1642	1643	1644	1645	1646	1647	1648	1649	1650	1651	1652	1653	1654	1655	1656	1657	1658	1659	1660	1661	1662	1663	1664	1665	1666	1667	1668	1669	1670	1671	1672	1673	1674	1675	1676	1677	1678	1679	1680	1681	1682	1683	1684	1685	1686	1687	1688	1689	1690	1691	1692	1693	1694	1695	1696	1697	1698	1699	1700	1701	1702	1703	1704	1705	1706	1707	1708	1709	1710	1711	1712	1713	1714	1715	1716	1717	1718	1719	1720	1721	1722	1723	1724	1725	1726	1727	1728	1729	1730	1731	1732	1733	1734	1735	1736	1737	1738	1739	1740	1741	1742	1743	1744	1745	1746	1747	1748	1749	1750	1751	1752	1753	1754	1755	1756	1757	1758	1759	1760	1761	1762	1763	1764	1765	1766	1767	1768	1769	1770	1771	1772	1773	1774	1775	1776	1777	1778	1779	1780	1781	1782	1783	1784	1785	1786	1787	1788	1789	1790	1791	1792	1793	1794	1795	1796	1797	1798

Qy	1533	PhMetGlyValAlaGlnIleLeuLeuAspGluLeuSerAsnMetValIleGly	1555
Db	621	TTTATGGGTGGCTCAGATCTTGTGTGAAGAACTCGACCTGTCCAGCATGGTGCATCGGA	680
Qy	1553	TripHelysLeuPheProProSerSerLeuValAspProThrSerAlaProLeuThrArg	1572
Db	681	TGTACAAATGTTCACCGTCTCACTGTGGATCCCACTCACTCCCTCACCCGG	740
Qy	1573	ArgAlaSerGlnSerSerLeuGluSerSerThrGlyProSerTyrSerArgSer	1590
Db	741	CGGGCTCCAGTCACTCTGGAAGATTCACTGGGCTCCCTGTATTGATCA	794
RESULT 6			
ABL12715	ABL12715 standard; cDNA; 7395 BP.		
XX	AC	ABL12715;	
XX	AC	26-MAR-2002 (first entry)	
DT	XX	Drosophila melanogaster expressed polynucleotide SEQ ID NO 32627.	
DE	XX	Drosophila; developmental biology; cell signalling; insecticide;	
KW	KM	pharmaceutical; gene; ss.	
XX	OS	Drosophila melanogaster.	
XX	PN	WO200171042-A2.	
XX	PD	27-SEP-2001.	
XX	PP	23-MAR-2001; 2001WO-US09231.	
XX	PR	23-MAR-2000; 2000US-191637P.	
PR	PR	11-JUL-2000; 2000US-0614150.	
XX	PA	(PEKE) PE CORP NY.	
XX	FI	Venter JC, Adams M, Li PWD, Myers EW;	
XX	DR	WPI; 2001-656860/75.	
DR	DR	P-PSDB; ABB68612.	
PT	PT	New isolated nucleic acid detection reagent for detecting 1000 or more	
PT	PT	genes from Drosophila and for elucidating cell signalling and cell-cell	
XX	XX	interactions -	
PS	PS	Claim 1; SEQ ID NO 32627; 21pp + Sequence Listing; English.	
CC	CC	The invention relates to an isolated nucleic acid detection reagent	
CC	CC	capable of detecting 1000 or more genes from Drosophila. The invention is	
CC	CC	useful in developmental biology and in elucidating cell signalling and	
CC	CC	cell-cell interactions in higher eukaryotes for the development of	
CC	CC	insecticides, therapeutics and pharmaceutical drugs. The invention	
CC	CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA	
CC	CC	sequences (ABL01840-ABL16175) and the encoded proteins	
CC	CC	(ABE57737-ABE72072).	
CC	CC	The sequence data for this patent did not form part of the printed	
CC	CC	specification, but was obtained in electronic format directly from WIPO	
CC	CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX	XX	Sequence 7395 BP; 1694 A; 2319 G; 2135 G; 1247 T; 0 other;	
Alignment Scores:			
Pred. No.:	5.32e-39	Length:	7395
Score:	992.00	Matches:	514
Percent Similarity:	28.11%	Conservative:	201
Best Local Similarity:	20.20%	Mismatches:	495
Query Match:	11.97%	Indels:	1336
DB:	23	Gaps:	72
US-09-617-099B-1 (1-1590) x ABL12715 (1-7395)			

Db 1959 CGTCCACGCCCGCCCTCCGCCACCGGCCCGTCCACGGCTCGGGCTTCGGGTTCC 2018
QY 810 ----- 810
Db 2019 CGCCCTGCGGCTTGAACGACAAACGACGGGGCCACCATACACACCGCCGCA 2078
QY 810 ----- 810
Db 2079 CCACTACACCCGCCACGAGTCCGCTTCGGATCCGACACGAGCATCAGGGTCGGCCGC 2138
QY 810 ----- 810
Db 2139 CGCCACCTGCAAGGCCACGGGTGGCGGTGGATCGGCATCGGCATCGGCAGCGCTCCGGCAC 2198
QY 811 ----- 811
Db 2199 CACCAACAGCCATCCCATCGAGGACGGCTGCAGCTTGGCTACACACAGAA 2258
QY 827 LGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuProSerArgGluAspGI 847
Db 2259 CACCTTGCAGCTGATGTAACCTGTGTCTGCGCACCGGACTCTCCCTGCGCCAGAGCGG 2318
QY 847 YArgProArgAspProTyrValIleLysIleTyrPheLeuProAspArgSerAspLysAsnLys 867
Db 2319 AGCGGGCGCAATCCATATGCAAAAGTGTCTCTGCGGACCGGACTCTCCCTGCGCCAGAGCGG 2378
QY 867 sArgArgThrLysThrValIleLysThrLeuGluProLysTyrPheAsnGlnThrPheIleTyr 887
Db 2379 GCGGCAACGAACAGCGTGGGCCACCTCGCAACCCCGGATGGGGCGGAGACCTTTGTCTA 2438
QY 887 rSerProValHisArgArgLysPheArgGluArgMetLeuGluIleThrLeuTrpAspGI 907
Db 2439 CTGGGTCTGAGCGGTGGATCTCAACGGCCGACTGTCTGAGTGACGCTGTGGACTA 2498
QY 907 nAlaArgValArgGluGluSerGluPheLeuGlyGluIleLeuIleGluLeuLuth 927
Db 2499 T-----GTGCGGTACGGGGCCAAACGACTTCATCGCGAGGTGTATCATGATCTGGCGCA 2552
QY 927 rAlaLeuLeuAspAspLupProHisTyrPheLysLeuGlnThrHis----- 942
Db 2553 CCACATCTCTGGACGAGGCGGAGTGTATCAGCTGCAGCCCGCACCGACACCTCTCTA 2612
QY 943 -----AspValSerSerLeuProLeu----- 949
Db 2613 TCTTTAGTGACGAGGCGAGCATGTGACGCGCTGATGACTGACACCGGACAGATCATTT 2672
QY 949 ----- 949
Db 2673 ATCACCGCGAGCACCATGTGCGGCTTAGTGATTCGGACACCGTCCGACTGGACAT 2732
QY 950 -----ProArgProSerPr 954
Db 2733 CGATGAATGACTCCGGGGGCCAGCATCTCGTCCATGGGCGAGCTCAGCGAGTCCGCGCC 2792
QY 954 oTyrLeu-----ProArgArgGlnLeuHisGlyGluSerProThrArg 969
Db 2793 CTGTCTAGAGCTCGATCAACAGAGGTGCTCCGACGTGACATGTCGCCC----- 2844
QY 969 gLeuGlnArgSerLysArgIleSerAspSerGluValSerAspTyrAspCysGluAspGI 989
Db 2845 ----CAGGGCCGCAACGGGTGGCGGATGGTGGCGCGGACTACCGCACTGTATCTGG 2900
QY 989 yValGly-----ValValSerAspTyrArgHisAsnGlyAr 1001
Db 2901 CATTGCAAAAGTTACCAATCAGGCTCTGCGCACCGGCTACTATGTCGCGGAGGTGG 2960
QY 1001 gAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSer-----SerAsnHi 1019
Db 2961 TAATGTGTGGCTTCGCCATCGGTCCCGGTGGATGAGCTTCAGCGAGCGGAGTCACTC 3020
QY 1019 sCysSerProSerGlySerProHisArgValAspValIleGlyArgThr----- 1035

Db 3021 GGCGGCACCCAGCGAGCGTTTACCACAGCAGCGGAGTAGGAGGAGCGGCTTCGGCGGAC 3080
QY 1036 -----ArgSerTrpSerProSer-----AlaProProProGlnAr 1047
Db 3081 CGCTTAAGGCTATCGGAGCACCGAGTCCCGGAGACGGGGCTCTCTTTCGCCACCGCCAGCCG 3140
QY 1047 g-----AsnValGI 1050
Db 3141 CTACATAGACTATCCAGTGTCTCCAGTACACGGCTCTCTCGCTTACGCCCCCATCGGTCTA 3200
QY 1050 uGlnGly----- 1052
Db 3201 CCAGGACCGGGTGTGTGTCTCTCAGCGCGGCGCATCTGCTCGCAGACAGCGATTCCA 3260
QY 1052 ----- 1052
Db 3261 GTGCGGATCGGCCACAGCCACCGCCACGGGATCTCCAAAGAGAGGAGCGACTACACAGGT 3320
QY 1052 ----- 1052
Db 3321 GCGCAGACCTCACGCGATGCGATGCTGCGAGATCGACTGGGTTCAGGACTTTCAGCAGCG 3380
QY 1053 -----HisArgGlyTh 1056
Db 3381 GCTGGCTCGGGTGTGTCTCGGGCGGCGCACCGCACACGGCAGCGCCACCCAGCGCCAC 3440
QY 1056 r-----ArgAlaThrGly-----HisTyrAsnThrIleSerArgMetAs 1069
Db 3441 CTACCGCAGCACCGGAATGGCGGTGGAGCGCCACTACCGGCTGTCTCGGACAGCGA 3500
QY 1069 p-----ArgHisArgValMetAspAsp----- 1076
Db 3501 CCTGCACTCTATGGACGCGAGTGCAGCGACCAAGGCACTCGTTATCCCGCGCAGAGCTT 3560
QY 1076 ----- 1076
Db 3561 TATGGCGAGTTCGGCGACTCGGACATGGATGCGTGTGTGAGCGTTCAGCGCTTCAGCGCCTT 3620
QY 1076 ----- 1076
Db 3621 CTCCAGCAGTTCGAGCGACCGCGCATGTCGCGCGGACTCAGCTTCCCGCGCACTGGCG 3680
QY 1076 ----- 1076
Db 3681 CAATCTCTTTGGGTCCGCAACAGCTTCTTCGAGAGGAGGAGCGCAACCCATTACCGG 3740
QY 1077 -----HisTyrSerSerAspArg----- 1082
Db 3741 ACTGCGGTGCACCACTCGGAGCC-CGGAAGGCGCAATACCATCGAGGTGGATGACTGCG 3799
QY 1082 ----- 1082
Db 3800 ACTATCTCGCGGTTCGGTTCGCCAGAGTGTGTGTCTGAGCAGCTGGAGCAACTGC 3859
QY 1083 -----AspArgAspCys----- 1086
Db 3860 AGCAGTCCCGCTCTTTCGGCAGCGAGTTCGCCACCCATTTCCGGCGGGAATTGGATAC 3919
QY 1086 ----- 1086
Db 3920 AGCCACCTCACCAAGTTTGTGTTTACGATGCCAATAGCGGCCAATTTGGTGGACAACTCT 3979
QY 1087 -----GluAlaAlaAspArgGlnPro----- 1093
Db 3980 AGCTTGAGCCCGCACATGAGGCGGAGGATGATGAGCGCACTCGATCCCATTTACATC 4039
QY 1093 ----- 1093
Db 4040 CACATTCCGCTCTTACTCGCATCGCCCATTTGTCACACAGGCAATTCGACATCGATA 4099
QY 1094 ----TyrHisArgSerArgSerThrGluGlnArgProLeuLeuGluArgThr----- 1110
Db 4100 TCGATACCATATATCCACCACCGAGCGTCCAAAGGACTGCAACCGAACTGACAGCCACCG 4159

```
QY 1110 ----- 1110
Db 4160 CCGCTGCTATCCGTCGCTGCTGCGAGCTTTGAGCAATTCAGGCGCATTCACGCG 4219
QY 1111 ----- 1111
Db 4220 CCATCAGCAATTTCTTCGCGAGCTCTCGCCCGCGGCTCCATCCACAGCCCGCAGACC 4279
QY 1121 rAsnLeu----- 1124
   :|||
   :|||
Db 4280 ACACCTCCACAGCGCCAGTTGCTGGTGGGTATGTAGTGACCAACCAGGACAGAGCT 4339
   :|||
QY 1124 tArgSerMetProSerLeuMetThrGlyArgSerAla----- 1140
   :|||
   :|||
Db 4340 GCAGCCACTGCAGAAATGGAAGTCAGCGGCTGCTATCGACGCGCACCCACCTTCACC 4399
   :|||
QY 1140 oAlaLeuSerArgSerHisProArg----- 1154
   :|||
   :|||
Db 4400 TGACCGCGCGGTCTGGCAATCGCGCTTGCCAGGATCCGCTGTCTGCGGATGCC 4459
   :|||
QY 1155 ----- 1158
   :|||
   :|||
Db 4460 ATCCGACCCACAGCATCAGCCGATGGGTATGGGAGCCACCTCTTGGATCTGGGA 4519
   :|||
QY 1158 ----- 1158
Db 4520 TGTGCGGAGTGTGTGATCAGCACTTCTGGTGGATTAAACGGGACCGCAACTGAATC 4579
   :|||
QY 1159 ----- 1167
   :|||
   :|||
Db 4580 TTGGTGTGTTCCCACTTACCATTGTTCACAGCCCACTCCAAATGCCATCGGCGTGCCA 4639
   :|||
QY 1167 rg----- 1174
   :|||
   :|||
Db 4640 GGGGCAAAATGGCAACTCTGCACCTTCCGCCACCAATCAGTGTACGCCCTTTCGCC 4699
   :|||
QY 1174 ----- 1174
Db 4700 AATTGAGCGAGATGCCCTCGTGGCAGCGTTCCTCCATTTGGAGGCCAAGCGCAACCA 4759
   :|||
QY 1175 ----- 1190
   :|||
   :|||
Db 4760 CTTGATCTCTGAAGAAACCGAGCTGGAGCGAAGCGACTCTTTCACGAGGACCAATCGC 4819
   :|||
QY 1190 rgGlnMetLysLeu----- 1194
   :|||
   :|||
Db 4820 GSTCGTAGACTACGATGATCTGCACAGAGAGCTGGGTGCGAGCTCGAATTCGTAGC 4879
   :|||
QY 1195 ----- 1199
   :|||
   :|||
Db 4880 AGGACGAGTGATGCTTCGTCTGGAGACTATCCGTATCCCATTCGAATCCGTATCCA 4939
   :|||
QY 1199 lnValala----- 1210
   :|||
   :|||
Db 4940 GCTACTCACCCACCGCGGCAATGTGCGAGCGTACGGATCGGAGACCAATATTCACAAA 4999
   :|||
QY 1210 spTYR----- 1215
   :|||
   :|||
Db 5000 GTTATATGGTGCAAAATGGATGCAATCATCCGCTGAGCCACTTCGCATTTCTCGTTAA 5059
   :|||
QY 1216 ----- 1219
   :|||
   :|||
Db 5060 CGGATGACAAGATAGTGACCAATTACCTACGACTCGGATGTGGCTGGAGCGAGCGGGG 5119
   :|||
QY 1219 ----- 1219
Db 5120 TGGCTCCGCTCTTTTCGGGAGCGCACCGCAACGCTGGACTGGGCGATCCAGGAATC 5179
   :|||
QY 1219 ----- 1219
Db 5180 ACATGCTGTGATTTGCAGCGAAACAGCCAAACAAAATTTGTACGACCGCGGCGACCGT 5239
   :|||
QY 1219 ----- 1219
Db 5240 ATTTGAGAGCCCGAAGAAATCTCGCATCCGCCAGTCCGCGAGATGCCAATGCCAATCTTT 5299
   :|||
QY 1220 ----- 1226
   :|||
   :|||
Db 5300 CACCAATGGAGCCGGATCGGGAATGGATGGCTTAATGCTGCGGCGAGCGGAATGCTGG 5359
   :|||
QY 1227 ----- 1234
   :|||
   :|||
Db 5360 AACAGGCAACGGAATGGTAGCGGTGTGCACACACACTTACCAGTACACACAATAGCA 5419
   :|||
QY 1234 spSerAspValSerAspValSerAlaValSerArgThrSerSer----- 1248
   :|||
   :|||
Db 5420 ACCACCAATAGCCAAAGCAGTAGTGTAGTTAGTCCAAAGTCAACAATCAACAATCTC 5479
   :|||
QY 1248 ----- 1248
Db 5480 AACAAACAACAATTAACAACAATCAGTTAGGCAAAAGGAAAGAGGACCAACCG 5539
   :|||
QY 1248 ----- 1248
Db 5540 CCATTAGCGGATCGATGATGCGATCGGAGCAGCTGACGAAATCTAGTAGTGGAGCGGTG 5599
   :|||
QY 1249 ----- 1252
   :|||
   :|||
Db 5600 GTGCAACATCTGTCGCCGTGCTGGGCATCTTGTGTCGGTGGTACTTAATGCTAATAATG 5659
   :|||
QY 1252 heSerSerThrSerTyrMetSerValGlnSerGluArgProArg----- 1266
   :|||
   :|||
Db 5660 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5719
   :|||
QY 1267 ----- 1279
   :|||
   :|||
Db 5720 ATGGATCACCCACAATCTCAACAACAATAAAGCTGCCCATCATCAACACCAACAC 5779
   :|||
QY 1279 lnAsnArgGlnMetGlyValSerGlyLysAsn----- 1289
   :|||
   :|||
Db 5780 AAAACACGAACACCAATAAATGACGCCAACCACTGACGCTGCGCCACACACAACAAT 5839
   :|||
QY 1289 ----- 1289
Db 5840 CACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 5899
   :|||
QY 1290 ----- 1298
   :|||
   :|||
Db 5900 CAACCAATTAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 5959
   :|||
QY 1298 ----- 1298
Db 5960 CAACCGCAAAATGCAACAATAATGCTACAACAACCTGCCCAAAATCCAAACAACAACA 6019
   :|||
QY 1298 ----- 1298
Db 6020 ATAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 6079
   :|||
QY 1298 ----- 1298
Db 6080 CGGATGCTGACGCGAGCGCTGATGCTCGCTGGAGCGCATCGACTGGGATGCCATCGATG 6139
   :|||
QY 1299 ----- 1307
   :|||
   :|||
Db 6140 CTATGCTGGAGCAGCACTTTAGCGAGTACGACAAGGACAAGACCGCGCGGAGCGGTG 6199
   :|||
QY 1307 spGlySerGlnSer----- 1320
   :|||
   :|||
Db 6200 GTGGCTCCAGTACGCGAGCGGTGCGAGTGGCGGAGTAAGTGAATGAGTGGAGCTCTCG 6259
   :|||
QY 1320 hrSerGlyLysLysArgArgSerSerIleGlyAlaLysMetValalalleValGlyLeu 1340
   :|||
   :|||
Db 6260 ACACCGCAACGATCGGAAGAGGCGCGCGC----- 6301
   :|||
QY 1340 erArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyLysLysLeuA 1360
   :|||
```

Db 6302 AGCGATCCCGAAGAGGAGCGAGCGCATGGCAAAACGATATGGGCTTTGGAAAGAGGGTA 6361
Qy 1360 rSer-----ThrValGlnArgSerThrGluThr----- 1369
Db 6362 AGAACTCGTTACGGTCACCGCAGCGAGGAAGTGTGCCCGCGATATAACCGGGGAAC 6421
Qy 1369 ----- 1369
Db 6422 TGGAAACGGTGGCGGATTTGGCGTCAGTCTCGCGCGGTGGCGGTGGCGCGGTGGT 6481
Qy 1369 ----- 1369
Db 6482 CAGGTGTGTGGCGGGGGCGGCTGTCCCGCGGCTCGTCTCGAGGTGGAGCCCATCG 6541
Qy 1370 -----GlyLeuAlaValGluMetArgAsnTrpMetThr- 1380
Db 6542 AGCAGTTCTTGGCGATGTGTGCCCGCGGGAAGCTTAAGCTAGACTTGGGTTCAGC 6601
Qy 1380 ----- 1380
Db 6602 TAGCGCTCTTGTGTAAAGCAAGCAAGCAATACCAAGCACAGTTCGCAATAGCAGCA 6661
Qy 1381 -----A 1381
Db 6662 GCTCGACCAAGTCGGCTGTGTCAACCAATATCACACTACCAACACACATCCATACCAC 6721
Qy 1381 rGlnAlaSerArgGluSer----- 1387
Db 6722 AGCGGAGTCGCGCGGCTCGCGGCGGAGCATCCAGCGGTGGTGGAGGTACCGCGGTGC 6781
Qy 1388 -----ThrAspGlySerMetAsnSerTyr---SerSerGluGlyAsnL 1401
Db 6782 CTCGGGTGACAGTTACGTACGCGCGGAGCAATCTCATCATCATAGCTCCGAGGCTCAT 6841
Qy 1401 euIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyL 1421
Db 6842 CATTCCTCCCTTCGTTGGC-----AACGATGAGGCCCTCAATGATGCTGTGATGCC 6895
Qy 1421 euGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIleG 1441
Db 6896 TGGGCCCTGTGCTCACTGTTGTCGCGCAGGTGCTGGGAGTCTCCCTCGTGGCGACATCC 6955
Qy 1441 lnValGlyMetMetAspLysGlyGlnLeuValGluLeuIleAlaArgGlyL 1461
Db 6956 AGCTATCGCTGTGCCACCAAAAGGCTGTCTGGAGGTGGAGGTCCATACCGGGCCAGGGCT 7015
Qy 1461 euValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeuA 1481
Db 7016 TACAGCAAAAGCCCAATCAAAATGTTGCTGCTCCATATGTTAAGTGTACCTAGTGT 7075
Qy 1481 spAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeuGluProL 1501
Db 7076 CGGAAAGCGCTGTGTGGCAAGATGAAACTTCTCGGCTCGAGCAGCCCTGGATCCAC 7135
Qy 1501 euTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleLev 1521
Db 7136 TGATTCAGCAGCAGTGTGTTATTCAGCAGTCTTACACCGGTTCGATCTACAGATACCG 7195
Qy 1521 alTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeuL 1541
Db 7196 TGTGGGCGACTATGGCGCATCGAGAGAGGTGTTATGGGCGTGGCGCAGATATGC 7255
Qy 1541 euAspGluLeuGluLeuSerAsnMetValIleGlyTrpPheLysLeuPheProSerS 1561
Db 7256 TCGACGATCTGAATCTGTGCAATATCGTATCGCTGGTCAAGCTCTTTGGCACCACCT 7315
Qy 1561 erLeuValAsp---ProThrSerAlaProLeu---ThrArgArgAlaSerGlnSerSerL 1579
Db 7316 CACTGGTCAGTGTGCCCACTAATATAGGCTTAGGCTTAGGCGCTCATCATAGCCTCAC 7375
Qy 1579 euGluSer 1581
|:|:|:|

Db 7376 TCGACTCA 7383
RESULT 7
ABAl1415
ID ABAl1415 standard; cDNA; 526 BP.
XX AC ABAl1415;
XX XX
DT 23-JAN-2002 (first entry)
XX XX
DE Human nervous system related polynucleotide SEQ ID NO 422.
XX XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisklering; antianaemic; antiarthritic; cancer;
KW antineumatic; hepatotropic; cerebroprotective; antinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.
XX XX
OS Homo sapiens.
XX XX
PN WO200159063-A2.
XX XX
PD 16-AUG-2001.
XX XX
PF 17-JAN-2001; 2001WO-US01334.
XX XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0248523.
PR 08-NOV-2000; 2000US-0248524.
PR 08-NOV-2000; 2000US-0248525.
PR 08-NOV-2000; 2000US-0248526.
PR 08-NOV-2000; 2000US-0248527.
PR 08-NOV-2000; 2000US-0248528.
PR 08-NOV-2000; 2000US-0248532.
PR 08-NOV-2000; 2000US-0248609.
PR 08-NOV-2000; 2000US-0248610.
PR 08-NOV-2000; 2000US-0248611.
PR 08-NOV-2000; 2000US-0248613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-541565/60.
XX P-PSDB; ABB15089.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX Claim 1; SEQ ID NO 422; 1701bp + Sequence Listing; English.
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 526 BP; 151 A; 134 C; 132 G; 103 T; 6 other;

Alignment Scores:
Pred. No.: 4,34e-21 Length: 526
Score: 598.50 Matches: 143
Percent Similarity: 43.57% Conservative: 6
Best Local Similarity: 41.81% Mismatches: 16
Query Match: 7.22% Indels: 177
DB: 22 Gaps: 3

US-09-617-099B-1 (1-1590) x ABA11415 (1-526)

QY 1068 MetAspArgHisArgValMetAspAspHisTyrSerSerAspArgAspArgCysGlu 1087
Db 3 ATGGACAGACATCGTGTGTCATGATGACCATATTATTCAGATAGACAGAGGATTGTGAA 62
QY 1088 AlalaAspArgGlnProTyrHisArgSerArgSerThrGluGlnArgProLeuLeuGlu 1107
Db 63 GCAGCAGATAGACAGACCATATCAGATCCAGATCCAGATCAACAGAACACCGCCTCTCCCTGAG 122
QY 1108 ArgThrThrThrArgSerArgSerSerGluArgProAspThrAsnLeuMetArgSerMet 1127
|||||

```
Db 123 CGGACCACCCCGCTCCAGATCCACTGACGCTCTGATACAAACCTCATGAGTGGATG 182
QY 1128 ProSerLeuMetThrGlyArgSerAlaProSerProAlaLeuSerArgSerHisPro 1147
Db 183 CCTTCATGATGACTGGAAGATCTGCCCTCCCTCTCACCTCCCTATCGAGGTCTCATPCT 242
QY 1148 ArgThrGlySerValGlnThrSerProSerSerThrProGlyThrGlyArgArgGlyArg 1167
Db 243 CGTACTGGGTCTCTCCAGACAGCCCATCAAGTACTCCAGTCCGACGACGAGGGGCCGA 302
QY 1168 GlnLeuProGlnLeuProProLysGlyThrLeuGluArgSerAlaMetAspIleGluGlu 1187
Db 303 CAGCTTCCACAGCTTCCACCAAGGGAACGTTGGATAGAAAGCA ----- 347
QY 1188 ArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaGlySerAspProArgLeu 1207
Db 347 ----- 347
QY 1208 GluGlnAspTyrHisSerLysTyrArgSerGlyTyrAspProHisArgGlyAlaAspThr 1227
Db 347 ----- 347
QY 1228 ValSerThrLysSerSerAspSerAspValSerAspValSerAlaValSerArgThrSer 1247
Db 347 ----- 347
QY 1248 SerAlaSerArgPheSerSerThrSerTyrMetSerValGlnSerGluArgProArgGly 1267
Db 347 ----- 347
QY 1268 AsnArgLysIleSerValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGly 1287
Db 347 ----- 347
QY 1288 LysAsnLeuThrLysSerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAsp 1307
Db 347 ----- 347
QY 1308 GlySerGlnSerAspThrAlaValAlaGlyAlaLeuGlyThrSerGlyLysLysArgArgSer 1327
Db 347 ----- 347
QY 1328 SerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSer 1347
Db 347 ----- 347
QY 1348 GlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerThr 1367
Db 348 -----GGAGGTAAAAAATAAGGAGCACTGTCCAAAGAGTACA 386
QY 1368 GluThrGlyLeuAlaVal-GluMetArgAsnTrp-----MetThrArgGlnAlaSerAr 1385
Db 387 GAAACAGGCGCTGGCCGTGGGAAATGAGGAACCTGGKATGGACTCGACAGGCGAAGCGGAG 446
QY 1385 gGluSerThrAspGly-SerMetAsnSerTyrSerSer---GluGlyAsnLeuIlePheP 1404
Db 447 AGTCTTACAGATGGTTAGCATGAACAGCTAACAGCTCCAGAGGAANTCTGATTTTNC 506
QY 1404 ro 1404
Db 507 CN 508
```

RESULT 8

```
ABL12714/c
ID ABL12714 standard; cDNA; 20604 BP.
XX
AC ABL12714;
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 32624.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
```

```
pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX P-PSDB; ABB68611.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 32624; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABBS73737-ABBY2072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 20604 BP; 5097 A; 4095 C; 5444 G; 5968 T; 0 other;

Alignment Scores:
Pred. No.: 3,55e-13 Length: 20604
Score: 470.50 Matches: 115
Percent Similarity: 57.42% Conservative: 32
Best Local Similarity: 44.92% Mismatches: 62
Query Match: 5.68% Indels: 48
DB: 23 Gaps: 5

US-09-617-099B-1 (1-1590) x ABL12714 (1-20604)
QY 1373 ValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerThrAspGlySerMet 1392
Db 1779 ATTGATGAGAGATGATACAAAGAAATATCCACACGACGCTGTACTAAAGTCTTCT 1720
QY 1393 AsnSerTyrSerSerGluGlyAsnLeuIlePheProGlyValArg-----Leu 1408
Db 1719 TTACTGTATCCCCACCCGACCCATTCACCAATTCCTCATAGATTCCTCCCTTCGTTG 1660
QY 1409 AlaSerAspSerGlnPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGly 1428
Db 1659 CGCAACGATGGAGCCCTCATGATGTCGTGGATGGCTGGCCCTGGTCACTGGTGGT 1600
QY 1429 ArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLys 1448
Db 1599 CGCCAGGTGCTGGAGCTCCCTCGCTGGCGGACATCCAGCTATCGTGTGCCACCAAAA 1540
QY 1449 GlyGlnLeuGluValGluIleIleArgAlaArgGlyLeuValVal----- 1463
Db 1539 GGCTGTCTGGAGTGGAGGTCTACGGGCCAGGGCTTACAGGTACGTAGGACGCAATTC 1480
QY 1464 -----LysProGlySer 1467
Db 1479 TATAGCATCTTATGATTTTCTTAACACTCTATCTGTGATCCCGACCAAAAGCCCAA-TCA 1421
```


Db 1890 GTGACGAGCGGAGGAGTCCCACTCTAGATCCCGACCACTAGGGGTAGATCT---CG 1946
Qy 355 userArgSpGluTyrGluArgGlnAArgGluGluGluTyrGlnAlaArgTyrArgSe 375
Db 1947 TTTCTAGAACACACGCGCGCGGCGAGTCCCGCTCTAGAACACCTGCCAGCGGAGATC 2006
Qy 375 rAspProAsnLeuAlaArgTyrProValLysProGln----- 387
Db 2007 A-----CGATCCAGAACTCCACGAGCGTAGGTCTCGGTCTAGAACACCGAGCGCG 2057
Qy 388 -----ProTyrGluGluGlnMetArgGlnMetArgGlnHisAlaGluVa 399
Db 2058 GAGGGGAGGTCTCGGTCTAGAACACCTGCTAGGCGCAGATCTAGACCCGATCACCAGT 2117
Qy 399 l-----SerArgAlaArgHisGluArgArgHisSerAspValSerLeu----- 413
Db 2118 ACGACGCGAGGTCTCGGTAGATCACCAGCCAGGAGAGTGGCAGGTCAAGCTCTAGAAC 2177
Qy 414 -----AlaAsnAlaGluLeuGluAspSe 421
Db 2178 CCCAGCTAGACGTGGCGCTCACGCTCCAGAACCCCGACGAGCGTGGCGCGCTCACGCTC 2237
Qy 421 rArgIleSerLeuLeuArgMetAspArg-----ProSerArg----- 433
Db 2238 TAGAACCCCGCTAGACGCGAGTGGTCTCAGCTCCAGAACACCGACGAGGAGGAG 2297
Qy 434 -----GlnArgSerValSerGluArgAlaAlaMetGluAsnGl 447
Db 2298 GTCTCGGTCTAGGACACCAAGCAGGAGAGATCCCGCAGTAGAAGCTTAGTAGACGTGG 2357
Qy 447 nArgSerTyrSerMetGluArgThrArgGluAlaGlnGlyGlnSerSerTyrProGlnAr 467
Db 2358 AAGATCTCACTCT-----AGAACACCTCAAGAGAGGCGAGATCTGGCTCATCTTCAGA 2411
Qy 467 gThrSerAsnHisSerProThrProArgSerProIleProLeuAspArgProAs 487
Db 2412 CGCGAATAACAAATCCAGAACATCTCAAGAGAGAGCAGG---TCCAATTCAGCCCGAGA 2468
Qy 487 pMetArgArgAlaAsp-----SerLeuArgLysGlnHisHisLeuAspProSerSerAlaVa 506
Db 2469 AATGAAGAAATCTCGCATTTCTTCAAGCGGAGCAGGTCTCTCTTCCACCGCTCCAA 2528
Qy 506 lArgLysThrLys-----ArgGl 512
Db 2529 AGCAAAATCTCGCTGTCTTTGAGCGCGAGCCTTTTCAGGGTCTTCCCATCCCTAAGCA 2588
Qy 512 uLysMetGluThrMetLeuArgAsnAspSerLeuSerSerAspGln----- 527
Db 2589 AAAGTCACAGACACACCCAGCGCGAGTCGCTCTGGATCTTCCCAACCTAAAGCTAAATC 2648
Qy 528 -----SerGluSerValArgProProProArg----- 537
Db 2649 TAGAACGCCACCCAGACGAGTCGCTCAGTTCTTCTCCGCCACCTAAACAGAAATCTAA 2708
Qy 538 -----ProHis-----LysSerLysLysGlyGl 545
Db 2709 GACACCATCAAGACAAAGTCATTCAGTTCTCTCTCATCTCTAAAGTAAATCTGGAA 2768
Qy 545 yLysMetArgGlnValSerLeuSerSerGlu---GluGluLeuAlaSerThrProGl 564
Db 2769 ACCACCGAGGCUAGGGTCTCAATACAGATCCCGAGCCCAATGAGCAATCTGTAAACGCCACA 2828
Qy 564 uTyrThrSerCys-----AspAspValGluLeuGluSerGluSer----- 577
Db 2829 GAGACGGAGCTGTTTGAATCATCCTGACCTGAGTTGAAATCTAGGACCCCTTCTAG 2888
Qy 577 ----- 577
Db 2889 ACATAGCTGCTCAGGGTCTCTCTCTCTAGAGTGAAATCTAGCACACCTCCCGACAGAG 2948
Qy 578 -----ValSerGluLysGlyAs 583
Db 2949 CCCATCTAGGTATCATCTCCCAACCCCAAGTGAAGGCAATATATATACCAAGACAAG 3008

Qy 583 pSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSer----- 598
Db 3009 AAGCCATTCCTGGCTCTCTTCTCAAGTCTCTAGTGGGTACGCTCGAGAAACAACCTCCACG 3068
Qy 598 ----- 598
Db 3069 GCGAAGCAGATCAGTATCTCTCCCTGCTCCAATGTGGAATCCAGATTGTGTCCAGATACAG 3128
Qy 599 -----AspSerAsnThrArgSerGluArgGlnLysLysArgMe 611
Db 3129 TCATTCCTGGGTCTCTCTCACCAGATACCAAGTGAACCTGAAACACCGCCAGACAAAG 3188
Qy 611 tTyrTyrGlyGlyHisSerLeuGluGluAspLeuGluTyrSerGluPro----- 627
Db 3189 TCATCAGGTCTATTTTCCATACCCCAAGTAAAGGCCCAAACTCCACCGGGGCCAAG 3248
Qy 628 -----GlnIleLysAspSerGlyValAspTh 636
Db 3249 TCCTTCGGATCAAGTCACCATGTCCCCAAGAGAGTCTTAAGACTCAGTGTCAAG 3308
Qy 636 rCys-----SerSerThrThrLeuAsnGluGl 645
Db 3309 TTGCGCTGGATCCCTCTCTCTGTGCGAGGTAAATCTAGCACACCCAGCGCGAGAG 3368
Qy 645 uHis----- 646
Db 3369 CTATTTTGGTGTCTCATCTCTGCAACTGAAGACAAATCTCAAACCTTCAACGACCAACAG 3428
Qy 647 -----SerHisSerAspLysHis----- 652
Db 3429 ATCTGATACTCAAGTCCAGAAAGTGAGACAGATCATTCAGATCACCATCTCTGCGAGAG 3488
Qy 653 -----ProValThr 656
Db 3489 CAAATCTCAAAACATCACCTAAGGAGGTGGTCCAGGTCTTCATCTCCAGTCACTGAGCT 3548
Qy 656 rGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuL 676
Db 3549 GGCATCCAGATCTCCAATAGACAAGATAGA-GGTGAGTCTCTCAGCGAGTCTTATGTATGA 3607
Qy 676 yAspGlySerValProArg----- 682
Db 3608 AATCTGGAATGTCTCTCTGAGCAGCAGGAGTTCGAGTCTGACTCTTCTCATATCTCTACAG 3667
Qy 683 --AspSerGlyAlaMetLeuGlyLysValValGly-----G 695
Db 3668 TGGACTCGAATTCCTCTCTGGGGCAGAGTAG-ATTGGAGACTGCTGAATCAAAAGAGAAA 3726
Qy 695 lLysMetThrGluSerGlyArgLeuCysAlaPheIleThrLys-ValLysLysGlySer 714
Db 3727 ATGGCTTACCCCTCAGGAGGA--TGCTACTGCATCACCCTCTAGACAGAAAGACAAA 3783
Qy 715 LeuAlaAspThrValGlyHisLeuArgPro----- 724
Db 3784 TTTAGTCCCTTTCAGTACAGGATAGCGCTTCACCTGCTTTCATGTTTCAAGACACACTT 3843
Qy 725 -----GlyAspGluValLeuGluTyrPasn 732
Db 3844 AGAACCCCGCCAGAGAAAGAGTGGTGTGGTCTCATCTCCAGAAACAAGAGACAAAT 3903
Qy 733 GlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleIleLeuGluSerLys 752
Db 3904 AGTCATTCCTACGTCAAGCCCAAGATGAAGAGTTAATGGAGGTGGTAGAAGTCTGAA 3963
Qy 753 ProGluProGlnValGluLeuValSerArgProIleGlyAspIleProArgIlePro 772
Db 3964 -----GAACCCGAGGCCAAATC-----CTGCT 3987
Qy 773 AspSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSer-----GlnLys 790
Db 3988 CATTTGCTTTCAGAACTTAAAGAAATGTCCACAGTAACTTTGAATCATCTCTCTGAAGTA 4047

ID AAL07156 standard; DNA; 24387 BP.
XX AAL07156;
AC
XX
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 9844.
XX
XX Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX Homo sapiens.
XX OS
XX WO200155320-A2.
XX
XX PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251938.


```
Qy 581 sGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSer----- 598
Db 13435 ACAAGAAGCCATTCTGGCTCTCTCTCTCAAGTCTTAGTAGGCTGAGTCGAGACAAC 13376
Qy 598 ----- 598
Db 13375 TCCACGGGAAGCAGATCAGTATCTCCCTGCTCCAAATGTGGAAATCCAGATTGTGGCAAG 13316
Qy 599 -----
Db 13315 ATACAGTCATTCTGGTCTCTCTCCACAGATACCAAAGTGAACCTGAACACCCGCCAAG 13256
Qy 609 sArgMetTyrTyrGlyGlyHisSerLeuGluGluAspLeuGluTrpSerGluPro----- 627
Db 13255 ACAAAAGTCACTCAGGGTCTATTTCACCATACCCCAAGTAAAGGCCCAACTCCACCGGG 13196
Qy 628 -----
Db 13195 GCCAAGTCTTCTGGATCAAAAGTCACCATGTCCCAAGAGAGTCTTAAGACTCACTAGT 13136
Qy 634 lAspThrCys-----
Db 13135 TCAAGTTCGCTCGATCCCTCTCTCTGTGCGAGGAGTAAATCTAGCACACCACCGAG 13076
Qy 643 nGluGluHis----- 646
Db 13075 CGAGAGCTATTTTGGTGTCTCATCTCTGCAACTGAAGAGCAAACTCAAACCTTCAACAGA 13016
Qy 647 -----
Db 13015 CCACAGATCTGATCTCAAGTCCAGAAAGTGAGACAGAGTCATTTCAGAAATCACCATTCTCT 12956
Qy 653 -----
Db 12955 GCAGAGCAAACTCAACATCACCTAAGGAGGTCCGGTCTTCTATCTCCAGTCCAC 12896
Qy 654 alThrTrpGlnProSerLysAspGlyAspArgLeuIleGlyArgIleLeuLeuLeuLysA 674
Db 12895 TGAGTGCATCCAGATCTCCAATAAGACAAGATAGA-GGTGAGTCTTCAGCGAGTCTTA 12837
Qy 674 rGlyLeuLysAspGlySerValProArg----- 682
Db 12836 TGTGAAATCTGGAATGCTCTCTGAGCAGAGAGAGGTTCCAGTCTGACTCTTCTTCATATC 12777
Qy 683 -----
Db 12776 CTACAGTCGACTCGAATCTCTCTTGGGCGAGAGTAG-ATTGGAGACTGCTGAATCAAAA 12718
Qy 695 -----GlyLysMetThrGluSerGlyArgLeuCysAlaPheIleThrLys-ValLysLys 712
Db 12717 GAGAAAATGGCCTTACCCCTCAGGAGGA---TGCTACTGCATCACTCTCTAGACAGAAA 12661
Qy 713 GlySerLeuAlaAspThrValGlyHisLeuArgPro----- 724
Db 12660 GACAAATTTAGTCCCTTCCAGTACAGGATAGGCTGAGTCTCACTGCTATTAACAGAC 12601
Qy 725 -----
Db 12600 ACACCTTAGAACCCCGCAAGAGAAAGAGTGTGTGGTCTATCTCCAGAAACAAGAG 12541
Qy 731 TrpAsnGlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnIleIleLeuGlu 750
Db 12540 CAAAATAGTGCAATGCTCCTACGTCAAGCCAAGATGAAGATTAAATGAGGTGTAGAGAG 12481
Qy 751 SerLysProGluProGlnValGluLeuValValSerArgProIleGlyAspIleProArg 770
Db 12480 TCTGAA-----
Qy 771 IleProAspSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSer----- 788
Db 12456 CTGTCTCATTTGCTTCAGAACTTAAGAAATGTCCACAAAGTAATTTGATATCATCTCTCT 12397
Qy 789 GlnLysMetAspArgProSerIleSerValThrSerProMetSerProGlyMetLeuArg 808
Db 12396 GAAAGTAGAAGAGGCTCTGTGTCTTACTCTTGATCAGAGCCAGTCACAGGCTTCT 12337
Qy 809 -----AspValProGlnPheLeuSerGlyGlnLeuSerIleLysLeuTrp----- 823
Db 12336 TTGGAAGCAGTAGAAGTCCCTTCAATGGCCTCATCT-----TGGGGT 12295
Qy 824 -----PheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 12294 GGGCCACATTTTCTCCAGAACATAAAGAACTG-----TCTAAC 12256
Qy 841 LeuProSerArgGluAsp-----GlyArgPro-----ArgAsnProTyrValLys 855
Db 12255 TCCCACTCAGGAGAACAGACTTTGGATCACCTTTAGAAATTTAGAAACTCAGGC----- 12202
Qy 856 IleTyrPheLeuProAspArgSerAspLysAsnLysArgArgThrIleValLysLys 875
Db 12201 -----CCACTTGGTACAGAAATGAATACTGATTTCTTCTGAGGTTAAAGAA 12154
Qy 876 ThrLeuGluProLysTrpAsnGlnThrPheIleTyrSerProValHisArgGluPhe 895
Db 12153 GATTTG-----AATGGACCGTTTCTTAA-TCACTCGAACAACAGATCCATCTCT 12107
Qy 896 Arg---GluArgMetLeuGluIleThrLeuTrpAspGln----- 907
Db 12106 AGATGAAAGAACATCGACAAGATCTCTGACACAGCAGGTTCTGAGTTATCCCCAGA 12047
Qy 908 -----AlaArgValArgGlu-GluGluSerGluPheLeuGlyGluIleLeuLeuLe 925
Db 12046 TGCAGTGGAAAAGCAGGATGTTCTTCAATCAG-----AGCAT 12008
Qy 925 uGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGlnThrHisAspValse 945
Db 12007 CTCTTCACTGTGTCTGATGCTGTACCCAGAACACCCCTCGAGAGAAAGAGTAGTTCTG 11948
Qy 945 rSerLeuPro-----LeuProArg---ProSerProTyrLeuProArgAr 959
Db 11947 ATCTTCTCTGAAATGAAGATGTTTACCAGAACTCCATCA-----AGAG 11900
Qy 959 gGlnLeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspse 979
Db 11899 AAGCAGGTCTGGGTCTTCTCCAGACTTAGA----- 11869
Qy 979 rGluValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAs 999
Db 11868 -----GATGGTCTGGGACTCCCTCG-----AGCCACAG 11840
Qy 999 nGlyArgAspLeuGlnSerSerThrLeuSerValProGlu---GlnValMetSerSerAs 1018
Db 11839 CCTGTCTGGGTCTCTCTGGAATGAAGATATACCTAGAACGCCATCTAGAGGGAGAAG 11780
Qy 1018 nHisCysSerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTr 1038
Db 11779 CGAATGTGATTTCTCCCAAGAACCGAAAGCTTTGCTCTCAGACTCTTAGCCGAGAGTCG 11720
Qy 1038 pSerProSerAlaPro-----ProProGlnArgAsnValGluGl 1051
Db 11719 TTCTCCATCATCCCGAGAGCTCAACAACAAGTGTCTTACCCCGCAGAGAGAA----- 11668
Qy 1051 nGlyHisArgGlyThrArgAlaThrGlyHisTyrAsnThrIle----- 1065
Db 11667 -----AGAAGCGGGTCAGAAATCATCATGTTGATCAGAAACTGTGCTCGGACTCCCTCGG 11612
Qy 1066 -----SerArgMetAspArgHisArgValMetAsp----- 1075
Db 11611 GCAGAGAAAGTCTGTGGGATCTCTCAAGAACTTGTGTGAACCCAGTGCATCCCTCA 11552
Qy 1076 -AspHisTyrSerSerAspArgAspArgCysGluAlaAlaAspArgGlnProTyr-- 1094
Db 11551 GGAAAGAGTGTGAGTCAGACTCTTCTCCAGATTTCTTAAGCCCAAGACACGACCCACTCG 11492
Qy 1095 -HisArgSerArgSerThrGluGlnArgProLeuLeuGluArg-----ThrThrTh 1111
```

```
Db 11491 GCAGAGGAGTCGGTCTGGATCATCTCCAGAGGTTGACAGCAAAATCTCGACTATCCCTCG 11432
Qy 1111 rArgSerArgSerSerGluArgProAspThr---AsnLeuMetArgSerMetProSerLe 1130
Db 11431 GCGCAGTAGGTCGTCTCCCTCCAGAGTGAAGTAAGCAAGAGCAGACCCAGGGC 11372
Qy 1130 uMetThrGlyArgSerAlaProProSerPro-----AlaLeuSerArgSerHisProAr 1148
Db 11371 ACAGAGTGGTCTGTATCTCTCTGACCTAAAGCTCCAGCCCTCGGCCCTCCAG 11312
Qy 1148 g-----ThrGlyse 1151
Db 11311 ACAGACAGATCAGGTCATCAAGCAAGGAGCAGAGCCCTTCTCCTGAAGGAAGCAGCAG 11252
Qy 1151 rValGlnThrSerProSerSerThrPro-----GlyThrGlyArgArgGly---ArgG1 1168
Db 11251 TACAGAGTCCTCTCTGAACTCCGCCCAAAATCCAGAACTGCTCGCAGAGGTTCCAGTTC 11192
Qy 1168 nLeuProGln-----LeuProProLysGlyThrLeuGluArgSerAl 1182
Db 11191 ATCACCAGAGCCCAAGACCAAGTCTGTACACCACTCGA-----CGTGGAGTTC 11141
Qy 1182 aMetAspIleGluGluArgGlnMetLysLeuAsnLysTyrLysGlnValAlaG1 1202
Db 11140 TCGATCATCTCCGGAGCTAAACAAGAGGCGCAGACTGTCCCGTAGAAGCGCTCTGCTC 11081
Qy 1202 ySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrrAspPro-- 1221
Db 11080 ATCTCTCACCA-----GAAACTCGCTCTAGAACTCCCCCAAG 11045
Qy 1222 -HisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSerAspValSe 1241
Db 11044 GCACGGAGAGTCTCCTAGTCTCTCC---CCGAGGCGAGCGCGGAAATCGAGTCTTC 10988
Qy 1241 rAlaValSerArgThrSerSerAlaSerArgPheSerThrSerTyrMetSer---Va 1260
Db 10987 ACGCCGCGGCGCTCAGCTTCATCTCCAGCACTAAGCAACCTCAAGAGAGCGCGCTC 10928
Qy 1260 lGlnSerGluArgProArgGly-----AsnArgLysIleSerVa 1273
Db 10927 TCCTTCGCCAAGACCTCGTGGACTCCAGAGTCCCGTTCCTCCCTCAAGAGAGAGAAAAA 10868
Qy 1273 lPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSe 1293
Db 10867 AAGAACACCCGACGTCGAGATAGTCTGGATCTTCAGTCAACCTCTCGCGGAAGACA 10808
Qy 1293 rThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGly-----Se 1309
Db 10807 GCGGAGCCGGTCAAGTCCGGTACTTCGGCGGCGGAGGAGGCTCTGGTTATCATCTC 10748
Qy 1309 rGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgSerSerIl 1329
Db 10747 AAGTGCACCTGCGCGGAGGAAGTTCGGACCTCCTCTCGACGCCGAAGA----- 10696
Qy 1329 eGlyAlaLysMetValAlaIleValGlyLeuSerArg-----LysSerArgSerAl 1346
Db 10695 -----GGCGCTCTCGACACCCCAACCCGAGTCGGAACG 10661
Qy 1346 aSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArgSer----- 1361
Db 10660 TTCTCGCTCAGCAGCATCACCAGCCCGCGTGGAAAGCTCTAGATCTCGAGCTCTCCAGC 10601
Qy 1362 -ThrValGlnArgSer---ThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetTh 1380
Db 10600 CACTCAGCGCGCATCAGGTCAGACCAACCCCTG-----ATAAG 10562
Qy 1380 rArgGlnAlaSerArgGluSerThr 1388
Db 10561 CCGACGTAGGTCCAGATCTCGAACT 10537
RESULT 12
AAV90372/c
```

AAV90372 standard; cDNA; 355 BP.

AAV90372;

15-FEB-1999 (first entry)

EST clone DM118.

Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
gene therapy; ss.

Homo sapiens.

WO9845436-A2.

15-OCT-1998.

10-APR-1998; 98WO-US069955.

10-APR-1997; 97US-0838821.

(GEMY) GENETICS INST INC.

Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
Racie LA, Spaulding V, Treacy M;

WPI; 1999-070077/06.

New polynucleotides encoding human secreted proteins - derived from
e.g. human blood, kidney, foetal lung, placenta, testes, brain,
ovary, pituitary, retina and colon cDNA libraries.

Claim 1; Page 519-520; 618pp; English.

The present sequence represents a human expressed sequence tag (EST).
The polynucleotide, which is a secreted EST, and the encoded protein
are predicted to have useful biological activities which would make
them suitable for treating, preventing or ameliorating medical
conditions in humans and animals, although no supporting data is
given. Suggested activities include nutritional activity, immune
stimulating or suppressing activity, haematopoiesis regulating
activity, tissue growth activity, activin/inhibin activity,
chemotactic/chemokinetic activity, haemostatic and thrombolytic
activity, receptor/ligand activity, anti-inflammatory activity,
cadherin/tumour invasion suppressor activity, tumour inhibition
activity. The polynucleotide may also be useful for gene therapy.

SQ Sequence 355 BP; 93 A; 82 C; 83 G; 97 T; 0 other;

Alignment Scores:

Pred. No.:	4.6e-10	Length:	355
Score:	368.00	Matches:	75
Percent Similarity:	96.34%	Conservative:	4
Best Local Similarity:	91.46%	Mismatches:	3
Query Match:	4.44%	Indels:	0
DB:	20	Gaps:	0

US-09-617-099B-1 (1-1590) x AAV90372 (1-355)

Qy 1272 SerValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThr 1291

Db 246 AGTGTCTTTACATCCAAATGCAAGCAGACAAATGGCATATCGGGAAGACATGACA 187

Qy 1292 LysSerThrSerIleSerGlyAspMetCysSerLeuGluLysAsnAspGlySerGlnSer 1311

Db 186 AAAAGCACCACATCATGTGGAGACATGTGCTCATCTGGAGAGATGATGGCAGCGCT 127

Qy 1312 AspThrAlaValGlyAlaLeuGlyThrSerGlyLysLysArgArgSerSerIleGlyAla 1331

Db 126 GACACTGCAGTGGGACCTTGGGACCACTGGGCAAGAGCGGCGCTCTAGCCTTGGTGGC 67

QY 1332 LysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeuSerGln 1351
 Db 66 AAAATGGTAGCTATCGTGTCTGTACGGAAAGTCGCAGTCTTCTCAGTCACGGGA 7
 QY 1352 ThrGlu 1353
 Db 6 CTCGAG 1
 RESULT 13
 AAV89630
 ID AAV89630 standard; cDNA; 355 BP.
 XX
 AC AAV89630;
 XX
 DT 15-FEB-1999 (first entry)
 XX
 DE EST clone CO618.
 XX
 KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
 KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
 KW gene therapy; ss.
 OS
 OS Homo sapiens.
 XX
 XX WO9845436-A2.
 PN
 PD 15-OCT-1998.
 XX
 PF 10-APR-1998; 98WO-US06955.
 XX
 PF 10-APR-1997; 97US-0838821.
 PR
 XX (GEMY) GENETICS INST INC.
 PA
 XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Racie LA, Spaulding V, Treacy M;
 XX
 XX WPI; 1999-070077/06.
 DR
 XX New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries.
 XX
 PS Claim 1; Page 279; 618pp; English.
 XX
 CC The present sequence represents a human expressed sequence tag (EST).
 CC The polynucleotide, which is a secreted EST, and the encoded protein
 CC are predicted to have useful biological activities which would make
 CC them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is
 CC given. Suggested activities include nutritional activity, immune
 CC stimulating or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The polynucleotide may also be useful for gene therapy.
 XX
 SQ Sequence 355 BP; 97 A; 83 C; 82 G; 93 T; 0 other;
 Alignment Scores:
 Pred. No.: 4.6e-10 Length: 355
 Score: 368.00 Matches: 75
 Percent Similarity: 96.34% Conservative: 4
 Best Local Similarity: 91.46% Mismatches: 3
 Query Match: 4.44% Indels: 0
 DB: 20 Gaps: 0
 US-09-617-099B-1 (1-1590) x AAV89630 (1-355)
 QY 1272 SerValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThr 1291
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 110 AGTGCTTTTACATCCAAAATGCAAGCAGACAAATGGGCATATCAGGAAGACATGACA 169
 QY 1292 LysSerThrSerIleSerGlyAspMetCysSerLeuGlyLysAsnAspGlySerGlnSer 1311
 Db 170 AAAAGCACCAGCATCAGTCGAGACATGTCTCACTCGAGAGAATATGCGACCCAGTCT 229
 QY 1312 AspThrAlaValGlyAlaLeuGlyThrSerGlyLysLeuArgSerIleGlyAla 1331
 Db 230 GACACTGCGAGTGGGCACCTTGGGCACCAAGTGGCAAAAGCGCGCTCTAGCCTTGGTGCC 289
 QY 1332 LysMetValAlaIleValGlyLeuSerArgLysSerArgSerAlaSerGlnLeuSerGln 1351
 Db 290 AAAATGGTAGCTATCGTGTCTGTACGGAAAGTCGCAGTCTTCTCAGTCACGGGA 349
 QY 1352 ThrGlu 1353
 Db 350 CTCGAG 355
 RESULT 14
 AAZ22301
 ID AAZ22301 standard; cDNA; 9551 BP.
 XX
 AC AAZ22301;
 XX
 DT 25-NOV-1999 (first entry)
 XX
 DE cDNA encoding a human trichohyalin (TRHY) protein.
 XX
 KW Human; trichohyalin; TRHY; protein; tissue structure; wound healing;
 KW terminally differentiating epidermal tissue; proteinaceous gel;
 KW breast implant; ss.
 XX
 OS Homo sapiens.
 XX
 PN US5958752-A.
 XX
 PD 28-SEP-1999.
 XX
 PF 14-FEB-1997; 97US-0800644.
 XX
 PF 30-APR-1993; 93US-0056200.
 XX
 PF (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Kim I, Chung S, Park S, Steinert PM, Lee S;
 XX WPI; 1999-561041/47.
 DR P-PSDB; AAY30795.
 XX
 CC Human trichohyalin useful for forming a proteinaceous gel that promotes
 CC wound healing -
 XX
 PS Claim 1; Fig 3A-W; 126pp; English.
 XX
 CC The present sequence encodes a human trichohyalin (TRHY) protein.
 CC The protein is found in terminally differentiating epidermal tissue,
 CC and is involved in forming the structural architecture of such
 CC tissue. The trichohyalin protein is useful for forming a
 CC proteinaceous gel which may then be used for healing wounds, or in
 CC breast implants.
 XX
 SQ Sequence 9551 BP; 2907 A; 1974 C; 2870 G; 1800 T; 0 other;
 Alignment Scores:
 Pred. No.: 3.5e-08 Length: 9551
 Score: 360.50 Matches: 324
 Percent Similarity: 33.02% Conservative: 242
 Best Local Similarity: 18.90% Mismatches: 568
 Query Match: 4.35% Indels: 586
 DB: 20 Gaps: 63
 US-09-617-099B-1 (1-1590) x AAZ22301 (1-9551)

Db 4830 GGAGCCGACAGCGGGGAACAACGGTTTCTCCCGAGGAGGAGGAGAA-----GGAGCA 4883
Qy 722 LeuArgPro-----GlyAspGluValLeuGluTrpAsnGlyArg----- 734
Db 4884 GCGCGCGCGCAGCGAGCGAGGAGGAGAGAGCTCCAGTCTCTGGAGGAAGAGGAGCA 4943
Qy 735 -----LeuLeuGlnGlyAla 739
Db 4944 GCTCCAGCGCGGAGCGTCCCAACAGCTCCAGGAGGAGGAGCGCTCCAGGAGGA 5003
Qy 740 ThrPheGluGluValTyraenilleLeuGluSerlyProGluProGlnValGluLeu 759
Db 5004 TCAGGAGAGGAGCGGACAGGA-----GCAGCGCGCGACCAAAAATGGAGGTG 5051
Qy 760 ValValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeu 779
Db 5052 GCAACTAGAGA----- 5063
Qy 780 GluSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerVal-Th 799
Db 5064 -----AGAAAGGAGAGAGAGCGCGCCACACAGCTGTACGC 5096
Qy 799 rSerProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSe 819
Db 5097 CAAGCCAGCCCTACAAGAGCAGCTGAGGAGAGAAACAGCAGCTGCTGCAGAGGAG----- 5151
Qy 819 rIleLysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLys 839
Db 5152 -----GAGGA 5156
Qy 839 sAspLeuProSerArgGluAspGlyA-ProArgAsnProTrpValLysIleTyPheLe 859
Db 5157 GGAGCTA----- 5163
Qy 859 uProAspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluPr 879
Db 5164 -----CAGAGAGAGAGCGCGAGAGAGAGAGCGCGCAAGAA----- 5199
Qy 879 oLysTrpAsnGlnThrPheIleTySerProValHisArgArgGluPheArgGluArgMe 899
Db 5200 -----CAGGAGAGACAAATACCGGAG----- 5220
Qy 899 tLeuGluIleThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGl 919
Db 5221 -----GAAGAGCAGCTGCAGCAGCAGGAGAGAGAGAGAGCTGTCTGAG 5258
Qy 919 yGluIleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTyT----- 937
Db 5259 AGAG-----GAACCGGAGAGAAAGAGAGCGCAGAGCGGGAGGCAATATCGGAA 5309
Qy 938 -----LysLeuGlnThrHisAspValSerSerLeuProLeuProArgProSerProTy 955
Db 5310 GGATAAGAGCTGCAGCAGAGGAA----- 5334
Qy 955 rLeuProArgArgGlnLeuHisGlyGluSerPro-----ThrArgArgLeuGlnArgSe 973
Db 5335 -----GAGCAGCTGCTGGGAGAGCAACCGGAGAGAGAGAGAGCGCGCAGAGCGGGA 5384
Qy 973 rLysArgIleSerAspSerGluValSerAspTyraSerAspCysGluAspGlyValGlyValva 993
Db 5385 GAAAAAATACCGCGAGGAGAG----- 5406
Qy 993 lSerAspTyraArgHisAsnGlyArgAspLeuGlnSerSerThrLeuSerValProGluGl 1013
Db 5407 -----GAGTTGACGAGGAG----- 5429
Qy 1013 nValMetSerSerAsnHisCysSerProSerGlySerProHisArgValAspValIleGl 1033
Db 5430 GCTGCTG-----AGAGAGGAGAGCGGAGGA 5453
Qy 1033 yArgThrArgSerTrpSerProSerAlaProProGlnArgAsnValGluGlnGlyHis 1053
Db 5454 GAGAGGCGCCAGGA-GTGGGAGAGCGAGTACCGCAAAAAAGACGAGCTGCAGCAGGAG 5512

Qy 1053 sArgGlyThr-----ArgAlaThrGlyHisTyraenThr-IleSerArgMetAspArgH 1071
Db 5513 AAGAGCAGCTGCTGAGAGAGAGAACCGGAGAGAAAGAGACTCCAGGAGCGGGAGAGCAAT 5572
Qy 1071 ieaArgValMetAspAspHisTyraSerSerAspArgAspArgCysGluAlaAlaAspA 1091
Db 5573 ATCGG-----GAGGAAGAGGAGCTGCAGCAGGAGGAAG 5605
Qy 1091 rGlnProTyraHisArgSerArgSerThrGluGlnArgProLeuLeuGluArgThrThrT 1111
Db 5606 AGCAGCTGCTGGGAGAGGAACCGGAGACGAGAGAGCGCCAGGAGCTGCAGAGCAATATC 5665
Qy 1111 hrArgSerArgSerSerGluArgProAspThrAsnLeuMetArgSerMetProSerLeuM 1131
Db 5666 GGAAGAGAGAGGAGCTGCAGCAGGAGGAAGAGCAGCTGCTGAGAGAGGA----- 5714
Qy 1131 eThrGlyArgSerAlaProProSerProAlaLeuSerArgSerHisProArgThrGlyS 1151
Db 5715 --ACCGAGAGAGAGAGCGCCGAGGAGCGGAGAGCAATGTCCGAGGAGAGGAGCTG 5772
Qy 1151 erValGlnThrSerProSerSerThrProGlyThrGly-ArgArgGlyArgGlnLeuPro 1170
Db 5773 CAGCAGGAGGAGAGCAGCTGCTGCAGAGAGGAACGGGAGAGAGAGGCGCCAG----- 5826
Qy 1171 GlnLeuProProLysGlyThrLeuGluArgSerAlaMetAspIleGluGluArgAsnArg 1190
Db 5827 -----GAGCTGGAGAGGCAATATCGGAGGAGAGAGAGCTTCACGCG 5868
Qy 1191 GlnMetLysLeuAsnLysTyraLysGlnValAlaGlySerAspProArgLeuGluGlnAsp 1210
Db 5869 CAGAAAGAGAAACAGCGATACCG-----GATGAGGAT 5901
Qy 1211 TyHisSerLysTyraArgSerGlyTyraAspProHis-ArgGlyAlaAspThrValSerTh 1230
Db 5902 CAGCGCAGTGATCTGAATCGCAGTGGAGACCAAGAAAAGAAAATGCAGTTCGTGATAAC 5961
Qy 1230 r-----LysSerSerAspSerAspValSerAspValSerAlaValSerArg-- 1245
Db 5962 AAGGTTTACTGCAAGGCGAGAGAGAAATGAACAGTTCGCGCAGTTGGAAGATTCCCGAGTG 6021
Qy 1246 ---ThrSerSerAlaSerArgPheSerSerThrSerTyraMetSerValGlnSerGlu-- 1263
Db 6022 CGCGACAGACAATCCAGCAAGATCTGCAGCAGCTGCTGGGTGAACAGCAGAGAGAGAGAT 6081
Qy 1264 -----ArgProArgGly---AsnArgLysIleSerVa 1273
Db 6082 CGTGAGCAAGAGAGGAGCGCTGCAGCAGCAGCGCAACAGGCAATTTCCAGAGAGAAACAG 6141
Qy 1273 lPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSe 1293
Db 6142 CTGAGCGAGAGAGCAAGCAAGGAAGCAAAAGGCGCGAGAGAGTCCCAAGAGGAAAG 6201
Qy 1293 rThr----- 1295
Db 6202 CAGTTGCTGAGAGAGAAAGAGAGAGAGAGAGAGCGCGCTCAGAGACAGACAGAAATTC 6261
Qy 1295 rIleSerGlyAspMetCys----- 1301
Db 6262 CGCGAGGAGGAACAGCTGCTCCAGGAAGGGAGGAACAGCGCTGCTCCCGCAGAGCGT 6321
Qy 1301 ----- 1301
Db 6322 GACAGAAAATTCGCGAGAGAGAACTGCTCCATCAGGAACAGGAGGAATAATTCCTCGAG 6381
Qy 1302 -----SerLeuGluLysAsnAspGlySe 1309
Db 6382 GAGGAACAGCGGCTCGCGAGGAACGGGAGAGAAAATTCCTTAAGGAGGAACAGCAGCTG 6441
Qy 1309 rGlnSerAspThrAlaValGlyAlaLeuGlyThr-----SerGlyLysLysArgAr 1326
Db 6442 CGCCTCGAGGAGCGGAGAGCACTCGTCAGGACCGCGACAGAAAATTCGCGAGGAGGA 6501

QY 1326 gSerSerIleGlyAlaIysMetValAlaIleValGlyLeuSerArgIysSerArgSerAl 1346
 Db 6502 CAGCAGCTGAGCGCCCAAGCGCTGCAGAAAATTCGTGAGAGGAACAGCAGGTGC 6561
 QY 1346 aSerGln-----LeuSerGlnTh 1352
 Db 6562 CGCCAGGACGAGAGAGAAAATTCCTGGAGGAGAACAGCAGCTGCGCCAGGAGCTCAC 6621
 QY 1352 rGluGlyGlyGlyLys----- 1358
 Db 6622 AGAAAATTCGCGAAGAGAGAACAGCTGCTCCAGAAAAGGAGAACAGCAGCTGCACCC 6681
 QY 1359 -----LeuArgSerThrVa 1363
 Db 6682 CAGACGCTGACAGAAAATTCCTGGAGGAGGAACAACAGCTGCGCCGACAGAGCTGAC 6741
 QY 1363 lGlnArgSerThrGluThrGlyLeuAlaVal-----GluMetArgAsnTrpMetThrAr 1381
 Db 6742 AGAAAATTCGCGAAGAGAACCTGCGCAGTCAGGAACACAGAGAAAATTCCTCGAGGAG 6801
 QY 1381 gGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsnLe 1401
 Db 6802 GAACAGCAGCTGACCGCCAGCAGCAACGCGCAGAGAAAATTCCTC-CAGGAGGAGACAGCAGCT 6860
 QY 1401 uillePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLe 1421
 Db 6861 GCGCGC-CAAGAGCGCGGCAACAGCGGCTCA----- 6893
 QY 1421 uGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly-AspIleG 1441
 Db 6894 -GGACCGTCACAGAAAATTCGCGAGGAGGAACAGCTGCGCCAGGAGGAGGAGGACAGC 6952
 QY 1441 lNValGlyMetMet-----AspLysGlyGlnLeuGluValGluIleArgAlaA 1459
 Db 6953 AGCTGAGCGCGCAAGAGCGCTGCAGAAAATTCGTTTAGAGGAACAGAAAGTGCAGCGC 7012
 QY 1459 rGlyLeuValValLys-----ProGlySerLysT 1469
 Db 7013 A-GGAACAGAGAGAAAATTCATGGAGGAGGAACAGCAGCTGCGCCGAGGAGGCGCAA 7071
 QY 1469 hrLeuProAlaProTyrValLysValTyrLeuLeuAspAsnGlyValCysIleAlaLysL 1489
 Db 7072 CAACAGCTGCGCA-----GGAGGACAGAAAATTCGCGGA 7107
 QY 1489 yslYstThrLys-ValAlaArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSer--- 1507
 Db 7108 GACCAACAGCTGCTCCAGGAGGAAGGAACAGCAGCTGCGCCAGCAGAGCTGACAGA 7167
 QY 1508 -----PheGluGluSerProGlnGlyArg 1515
 Db 7168 AAATTCCTCGAGGAGGAACGCGCAGCTGCGC 7197
 RESULT 15
 ID AAV21511
 XX AAV21511 standard; DNA; 6755 BP.
 AC AAV21511;
 XX 17-AUG-1998 (first entry)
 DT Staphylococcal bacteriocin BacRI operon.
 DE
 DE
 XX BacRI; bacteriocin; antimicrobial; antibacterial; antibiotic;
 KW Moraxella bovis; infectious bovine keratoconjunctivitis; cancer;
 KW therapy; ds.
 XX Staphylococcus aureus strain UT0007 (ATCC 55800).
 OS
 XX W09812319-A1.
 PN
 XX 26-MAR-1998.
 PD
 XX 18-SEP-1997; 97WO-US16758.
 PF

XX 17-SEP-1997; 97US-0931999.
 PR 19-SEP-1996; 96US-0710561.
 XX (UNIV) UNIV KANSAS STATE RES FOUND.
 PA Crupper SS, Iandolo JJ;
 PI WPI, 1998-230316/20.
 DR
 XX Therapeutic proteinaceous substances from *Staphylococcus aureus* -
 PT useful to inhibit growth of wide range of prokaryotic or eukaryotic
 PT cells, e.g. *Moraxella bovis* causing infectious bovine
 PT keratoconjunctivitis
 XX Claim 2; Page 19-23; 38pp; English.
 XX This polynucleotide comprises the bacteriocin BacRI operon of
 CC *Staphylococcus aureus* UT0007. The sequence of the BacRI operon
 CC was determined by N-terminal sequencing of purified BacRI peptide
 CC (see AAW54171), with back-translation and plasmid analysis. The
 CC BacRI operon includes the BacRI gene (see AAV21510), a homologue of
 CC the *cym* gene of the cytolytic operon of *Enterococcus faecalis*
 CC whose function is involved in the maturation of pre-cytolysin,
 CC an ATP-transporter gene, *bio1* and *bio2* genes related to
 CC lactococcal biosynthesis and modification, and a gene involved in
 CC immunity function. BacRI peptides can be produced by construction
 CC of an expression vector containing an oligonucleotide or operon
 CC coding for BacRI, and use of the vector to transform host cells for
 CC BacRI expression. The entire BacRI operon has been cloned into
 CC plasmid pUB110, and *Bacillus subtilis* transformants secreted the
 CC recombinant BacRI peptide into the medium. Bacteriocin BacRI is
 CC active against many Gram-positive and Gram-negative organisms such
 CC as *Bordetella bronchiseptica*, *Pasteurella multocida* and
 CC *Staphylococcus aureus*; *Moraxella bovis*, causing infectious bovine
 CC keratoconjunctivitis, is especially sensitive. BacRI can also be
 CC used as an anti-cancer agent.
 XX Sequence 6755 BP; 3903 A; 1252 C; 1600 G; 0 U; 0 other;
 SQ

Alignment Scores:

Pred. No.: 6,18e-06 Length: 6755
 Score: 311.00 Matches: 288
 Percent Similarity: 34.2% Conservative: 226
 Best Local Similarity: 19.2% Mismatches: 593
 Query Match: 3.75% Indels: 399
 DB: 19 Gaps: 53

US-09-617-099B-1 (1-1590) x AAV21511 (1-6755)

QY 8 ArgGlyArgProAlaProThrProAlaAlaSerGlnProProGln-----ProGlu 25
 Db 61 AGAGAAAGAGAGCGGAGAACACCCAGGGGGAGCGCCACCGGGGAGGAGAACCAACAAA 120
 QY 26 MetProAspLeuSerHisLeuThrGluGluGluArgLysIleIleLeuAlaValMetAsp 45
 Db 121 CAGCCCAACCAACCGCACGC-CCGCGAGGAGGAGAG-----AGC 158
 QY 46 ArgGlnLysLysGluGluLysGluGlnSerValLeuLysIleLys----- 61
 Db 159 AAAGGACACACAGAGAGAGAACCAACCAACAAAACGAGAAAGAACCAACCAAGAGAA 218
 QY 62 -----GluGluHisLysAlaGlnPro 68
 Db 219 GAAACGACCCAGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 278
 QY 69 ThrGlnTrpPheProPheSerGlyIleThrGluLeuValAsnValLeuGlnProGln 88
 Db 279 GCGAAGGGC-----GGAGACAAACGCGAAGCAACGCA 308
 QY 89 GlnLysGlnProAsnGluLysGluProGlnThrLysLeuHisGlnPheGluMetTyr 108
 Db 309 AAAAACAACAAACACGAGGAAAAAGGCGACCGCAACCGCGCAGAGAGAGAGAGAGAA 368

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 08:13:46 ; Search time 12923 Seconds
(without alignments)
5033.375 Million cell updates/sec

Title: US-09-617-099B-1
Perfect score: 8285
Sequence: 1 MSAPLGRGRPAPTPAASQP.....TTRASQSSLESSTGPSYSRS 1590

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgm2/1/USPTO.epool/US09617099/runat_14112003_191546_23681/app_query.fasta_1.1735
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -WAXLEN=200000000
-USER=US09617099.ecgm_1_18872@runat_14112003_191546_23681 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.mu.*

21: em.ov.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8285	100.0	4980	6	AX137063 Sequence
2	8285	100.0	4980	6	B63816 Protein Rim
3	7930	95.7	4800	10	AB021131 Mus muscu
4	7459.5	90.0	5550	10	AF199331 Rattus no
5	7438.5	89.8	5592	10	AF199330 Rattus no
6	7420.5	89.6	5640	10	AF199322 Rattus no
7	6856.5	82.8	5172	10	AF199324 Rattus no
8	6543	79.0	4965	10	AF199329 Rattus no
9	6522	78.7	5031	10	AF199323 Rattus no
10	6522	78.7	5031	10	AF199326 Rattus no
11	6522	78.7	5031	10	AF199327 Rattus no
12	6522	78.7	5031	10	AF199332 Rattus no
13	6504	78.5	5079	10	AF199325 Rattus no
14	6261.5	75.6	5063	10	AF548738 Rattus no
15	6212	75.0	4851	10	AF199328 Rattus no
16	5638	68.1	3854	9	AB018294 Homo sapi
17	5239.5	63.2	4456	9	BC043144 Homo sapi
18	4583	55.3	5079	9	AY190519 Homo sapi
19	4513.5	54.5	5841	10	AF199333 Rattus no
20	4489.5	54.2	5655	10	AF007836 Rattus no
21	4402	53.1	5325	9	AB051866 Homo sapi
22	2936.5	35.4	6395	9	AB002338 Human mRN
23	2936.5	35.4	6638	6	BD083751 Nucleic a
24	2936.5	35.4	6638	6	BD097398 Nucleic a
25	2936.5	35.4	6638	9	AB045726 Homo sapi
26	2424.5	29.3	2223	9	AF263306 Homo sapi
27	2410.5	29.1	2301	9	AF263305 Homo sapi
28	2340.5	28.2	2103	9	AF263307 Homo sapi
29	1912	23.1	1877	9	AF007156 Homo sapi
30	1834.5	22.1	1791	9	AF263308 Homo sapi
31	1821	22.0	1704	9	AF263309 Homo sapi
32	1763.5	21.3	1551	9	AF263310 Homo sapi
33	1612	19.5	254783	2	AC101935 Mus muscu
34	1580	19.1	139527	2	AC141382 Rattus no
35	1498	18.1	139398	9	AC025836 Homo sapi
36	1498	18.1	152336	9	AF004714 Homo sapi
37	1498	18.1	162489	9	AF001572 Homo sapi
38	1452.5	17.5	4692	3	AF257062 Caenorhab
39	1221.5	14.7	1963	10	AF199335 Rattus no
40	1042	12.6	7239	9	D87074 Human mRNA
41	1042	12.6	7325	6	AX281663 Sequence
42	1025.5	12.4	1424	9	BC003103 Homo sapi
43	1022	12.3	1605	10	AF199334 Rattus no
44	830	10.0	1163	10	AF310531 Mus muscu
45	828	10.0	11647	3	U41035 Caenorhabdi

ALIGNMENTS

RESULT 1

AX137063	AX137063	4980 bp	DNA	linear	PAT 30-MAY-2000
LOCUS	Sequence 2 from Patent EP1090986.				
DEFINITION	AX137063				
ACCESSION	AX137063.1	GI:14273409			
VERSION					
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1				
AUTHORS	Seino, S., Shibasaki, T. and Ozaki, N.				
TITLE	Protein rim 2				
JOURNAL	Patent: EP 1090986-A 2 11-APR-2001;				
FEATURES	JCR PHARMACEUTICALS CO., LTD. (JP) ; Seino, Susumu (JP)				
source	Location/Qualifiers				
	1. 4980				
	/organism="Mus musculus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:10090"				
BASE COUNT	1503 a 1161 c 1282 g 1034 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	3.04e-252	Length:	4980		
Score:	8285.00	Matches:	1590		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-617-099B-1 (1-1590) x AX137063 (1-4980)					
Qy	1 MetSerAlaProLeuGlyProAArgGlyArgProAlaProThrProAlaAalaSerGlnPro	20			
Db	32 ATGTGGCTGCTCGGCTCGGCGCGCGCGCGCTCCACCCCGCGGCTCTCACT	91			
Qy	21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysIlele	40			
Db	92 CTTCCGACGCCGAGATCGCGACCTCAGCCACCTCAGGAAGAGAGAGAAATCATC	151			
Qy	41 LeuAlaValMetAspArgGlnLysGluGluGluLysGluGlnSerValLeuLysIle	60			
Db	152 CTGGCTGTCTATGGATCGTCAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAG	211			
Qy	61 LysGluGluHisLysAlaGlnProThrGlnTTPheProPheSerGlyIleThrGluLeu	80			
Db	212 AAGAGAAACACAAAGCACACACGACAGTGGTTTCCCTTTAGTGGATCTACTGACTG	271			
Qy	81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys	100			
Db	272 GTAAATAACGTTCTCGAGCCCGAGCAAAACCAACCAATGAGAGAGAGAGAGAGAG	331			
Qy	101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGlnSerGln	120			
Db	332 CTGCACCAACAAATTTGAATGTATAGGAGCAAGTCAAGAGATGGAGAGAAATCGCAG	391			
Qy	121 GlnGlnGlnGluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe	140			
Db	392 CAGCAGCAGAGACAGAGAGGATGATCCCGAGCTGTGGCATCTGCCACAGACAAATTT	451			
Qy	141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly	160			
Db	452 GCAGATGATGATCGGCCATAATTTGTTCCATATGCCAAACCAAGTCTCTGCTCGATGGA	511			
Qy	161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln	180			
Db	512 GGTCTGAGTCTCTTACGCTCAACCAAGGTTATGTGGTGTGTAAATTTGCGCGCAACAA	571			
Qy	181 GlnGluIleLeuThrLysSerGlyAlaTTPheTyrAsnSerGlySerAsnThrLeuGln	200			
Db	572 CAAGAAATCCCTCACATAATCAGGAGCATGGTTTATATAGTGGGTCTAACACACTGCGAG	631			

Db 1712 TCCACACCTGAGTATCAAGCTGTGATGATGTGAGCTGGAAAGCGAGAGTGTGAGTGAG 1771
Qy 581 LysGlyAspSerGlnLysGlyAspGlyThrSerGlnGlnGlyValLeuSerAspSer 600
Db 1772 AAAGGGGACAGTCAAAAGGGGAAAGAAAGAACTAGTGAGCAGGAGGTTTGTGCGATCT 1831
Qy 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
Db 1832 AACACAGGCTCTGAGAGACAAAGAAAGAGATGACTATGTTGGTGGCCACTTTTGGAGAG 1891
Qy 621 AspLeuGlnTrpSerGluProGlnLysAspSerGlyValAspThrCysSerSerThr 640
Db 1892 GATTGGAAATGGTCTGAGCCTCAGATTAAAGACTCTGGGGTAGATACCTGTAGTAGCACA 1951
Qy 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
Db 1952 ACCCTTAAACGAGGAGCATAGCCNTAGTAGTAAGACACCCCTGTGACCTGGCAGCCATCCAAA 2011
Qy 661 AspGlyAspArgLeuLeuGlyArgGlnLeuLeuAsnLysAspGlySerVal 680
Db 2012 GATGGAGATGCCCTAATTTGTCGTATTTTAAATTAAGCGTTTAAAGATGGAGATGTA 2071
Qy 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
Db 2072 CCTCGAGACTCAGGAGCAATGCTGGGCTTAAAGGTTGTAGGAGGAAAGATGACTGAATCA 2131
Qy 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 2132 GGTGCGATTTGTGCAATTTATTAACCAATGTAAGAAAGGAAAGTTTGTAGCTGACTGTAGGA 2191
Qy 721 HisLeuArgProGlyAspGluValLeuGlnTrpAsnGlyArgLeuLeuGlnGlyValThr 740
Db 2192 CATCTTAGACGCGGTGATGAAGTCTTGAATGGAAATGGAGCGCTATTGCAAGGAGCCACA 2251
Qy 741 PheGluGluValTyrAsnIleLeuLeuSerLysProGluProGlnValGluLeuVal 760
Db 2252 TTTGAGGAAGTTTACAACTATTATCTAGATTCACAACTGAAACCTGAAACCAAGTTGAGCTTGT 2311
Qy 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
Db 2312 GTTTCAAGGCGCAATGAGATATTCCTAGATATCTGATAGACGATGCAACAACTGGAA 2371
Qy 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2372 TCCAGTTCTAGCTCATTTGAATCTCAAAAATGACCGCTCTCTATATCCGTTACCTCA 2431
Qy 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2432 CCCATGAGTCTGCGCATGCTGAGGGATGTCCTCGCAGTCTTATCTGGACAGCTTCAATA 2491
Qy 821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 2492 AAACATATGTTTGAACAGTTGGTCCACAGTTGATAGTTACAAATTTTGGGAGCAAGGAT 2551
Qy 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
Db 2552 CTCCCTTCCAGGAAGATGGAGGCCCAAGGAATCCTTATGTTAAGATTATCTTCCCTCCA 2611
Qy 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
Db 2612 GATAGAGTGATAAAATATAGAGAAGAACAAAACAGCTCAAGAAATCTTTGGAAACCCAAA 2671
Qy 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2672 TGGAAACAGACATTTCAATTTATCTCTGTCCTGCCAGAGAGATTCCTGTAAGCAATGCTG 2731
Qy 901 GluIleThrLeuTrpAspGlnAlaArgValArgGlnGluLeuSerGluPheLeuGlyGlu 920
Db 2732 GAAATTAACCTTTGGGATCAAGCTAGAGTTTCGAGAAGAAGAGAGCGAATTTCTTAGGAGAG 2791
Qy 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspGluProHisIleTrpTyrLysLeuGln 940

Db 2792 ATTTTAATTGAATTGGAAACACAGCTTTGCTAGATGATGAGCGGCACTGGTATAGCTGAG 2851
Qy 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
Db 2852 ACCCATGATGTCTCTCTCAATTCGCCACTCCCTGCGCCCTTCCCATATCTGCCCCGAGCAG 2911
Qy 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
Db 2912 CTCCATGGAGAGAGCCCAACGCCAGCGCTGCAAAAGTTCGAAGAGATTAAGTGCAGTGAA 2971
Qy 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
Db 2972 GTGTCTGACTACGACTCGCAGCATGGCTGGAGTAGTGTGCAGATTATCGACACAATGGC 3031
Qy 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
Db 3032 CGCGATCTTCAAAAGCTCCAGCTTGTGCGTCCAGAACACAGTCAATGTCATCAATCATTCG 3091
Qy 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
Db 3092 TCACCATCAGGTTCTCTCATCCAGTAGATGTTATAGGAAGACAGGTCAATGTCGCT 3151
Qy 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
Db 3152 AGTCCCCCTCTCTCAAGGAATGTGGAACAGGGGCACCGAGGACACGCTGCTACTGGC 3211
Qy 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
Db 3212 CATTAACAACAATTAGCCGATGGATAGACACCGTGTCAATGATGACCACTACTCTTCA 3271
Qy 1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
Db 3272 GATAGAGACAGGATTTGTGAACGACGAGATGACAGCCATATCACAGTCCAGATCAACA 3331
Qy 1101 GluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProAsp 1120
Db 3332 GAAACAACGGCTCTCTCTAGAGCGACACCAACCGCTCCAGATCTCTTGAACGCTCTGAT 3391
Qy 1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140
Db 3392 ACAAACTCATAGAGTCGATGCTTCAATTAATGACTGGAAGATCTGCCCTCTCTTCACT 3451
Qy 1141 AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrPro 1160
Db 3452 GCCTTATCGAGGTTCTACCTCTGATCCGGTCTGTCCAGACAAGCCCATCAAGTACTCCG 3511
Qy 1161 GlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProLysGlyThrLeuGluArg 1180
Db 3512 GGAACAGACGAAGGGGCCGACAGCTTCCACAGCTTCCACCAAGGGAAACATTGGAGAGA 3571
Qy 1181 SerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnVal 1200
Db 3572 AGTGCTATGGATATAGAGGAGAGAAATCGCAAAATGAACTTAACAAATACAAACAGTA 3631
Qy 1201 AlaGlySerAspProArgLeuGlnAspTyrHisSerLysTyrArgSerGlyTrpAsp 1220
Db 3632 GCCGATCAGACCCGACCTGGAGCAGATTAACATTCGAAGTATCTCCTCAGGATGGAT 3691
Qy 1221 ProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSerAspVal 1240
Db 3692 CCACATAGAGGGGCAGATATCTGTTTCCACTTAATCTCCGACAGCTGATGTAGTATGTA 3751
Qy 1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerVal 1260
Db 3752 TCTCGGTTTTCAAGGACTAGTAGTGTCTTCGTTTCAGCAGCACAAGTACATGTCGCTC 3811
Qy 1261 GlnSerGluArgProArgGlyAsnArgIleSerValPheThrSerLysMetGlnAsn 1280
Db 3812 CAATCAGAGCGCGAGAGGAAACAGGAAATCAGTGTCTTTTATCTTCAATTCCAAATGCAAAAC 3871
Qy 1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMet 1300
Db 3872 AGACAGATGGGCGTGTGCGGGAAGAACTTGACCAAAAGACCAAGCATCAGTGGAGACATG 3931

QY	1301	CysSerLeuGluValAsnAspGlySerGlnSerAspThrAlaValAlaLeuGlyThr	1320
Db	3932	TGCTCTCTGGAAGATGACCGCAGCCAGTCCGACACTGCAGTGGCGCCCTGGGTACC	3991
QY	1321	SerGlyLysLysArgSerSerIleGlyAlaLysMetValAlaLysValGlyLeuSer	1340
Db	3992	AGTGGCAAGACGGCGATCTAGCATTTGGGCGCAAAATGGTAGCTATTGTTGGTCTCA	4051
QY	1341	ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyGlyLysLysLeuArg	1360
Db	4052	CGGAAAGTCGAGTGCCTCTCAACTCAGCAAAACGGAAGGAGGAGTAAAGATACGG	4111
QY	1361	SerThrValGlnArgSerThrGlnThrGlyLeuAlaValGluMetArgAsnTropMetThr	1380
Db	4112	AGCATGTTTCAGAGAAGCGGAGACCGGCTAGCAGTGGAGTTCAGGAACCTGGATGCC	4171
QY	1381	ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsn	1400
Db	4172	CGCCAGGCCAGCGGGAATCCACAGATGGCAGCATGAACAGCTATAGCTCGAAGGAAT	4231
QY	1401	LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly	1420
Db	4232	CTGATCTTCCCTGGGTCGCGCTCGACAGCCAGTTCAGTCAATTCCTGGATGCC	4291
QY	1421	LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle	1440
Db	4292	CTGGGCCCCCTGCTCAGCTAGTGGAGCGCCAGACCCCTGGCTACTCTCTGCAATGGGTGACATT	4351
QY	1441	GlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleLeuArgAlaArgGly	1460
Db	4352	CAGGTGGGAATGATGATAAAGGACAGCTGGAGGTAGAATCATCCGGCGCGCGGC	4411
QY	1461	LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu	1480
Db	4412	CTTGTGTAAACCCAGGTTCCAGACACTGCCAGCACCGCTATGTCCTCAAGGTGATCTGTTA	4471
QY	1481	AspAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeuGluPro	1500
Db	4472	GCACCGAGTCTGATACCCAAAGAAACCAAGGTGGCGAGAAAGCCCTGGAGGCC	4531
QY	1501	LeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleIle	1520
Db	4532	CTGTACACGACGCTCTGCTTCGAGGAGAGCCCCAGGGAGGGGTGTACAGATCAT	4591
QY	1521	ValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu	1540
Db	4592	GTCTGGGAGATTATGGTCGTATGGATCAAAATCCTTTATGGGAGTGGCCAGATATC	4651
QY	1541	LeuAspGluLeuGluLeuSerSerMetValIleGlyTyrPheLysLeuPheProSer	1560
Db	4652	TTAGATGAACCTGAACCTATCCAACTATGGTATGGATGGTTCAAACTCTCTCTCTCC	4711
QY	1561	SerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGlu	1580
Db	4712	TCCTAGTAGATCCACCTCGCAGCTCTGACAGAGAGCTTCCCAATCGTCTCTGAA	4771
QY	1581	SerSerThrGlyProSerTyrSerArgSer	1590
Db	4772	AGTTCTACGGACCTTCTACTCTCGTTCA	4801
RESULT 2			
E63816			
LOCUS	E63816	4980 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Protein Rim2.		
ACCESSION	E63816		
VERSION	E63816.1	GI:22553654	
KEYWORDS	JP 2001103974-A/1.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 4980)		

AUTHORS Kiyono, S., Shibazaki, T. and Ozaki, N.
TITLE Protein Rim2
JOURNAL Patent: JP 2001103974-A 1 17-APR-2001;
SUSUMU KIYONO, NIHON CHEMICAL RESEARCH K K
COMMENT OS Mus musculus (mouse)
PN JP 2001103974-A/1
PD 17-APR-2001
PF 08-OCT-1999 JP 1999288372
PI SUSUMU KIYONO, TADAO SHIBAZAKI, NOBUAKI OZAKI
PC C12N15/09, C07K14/47, C07K16/18, C12Q1/68, G01N33/53//C12P21/08,
C12N15/00
CC

Key Location/Qualifiers.
FEATURES
source
1..4980
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 1503 a 1161 c 1282 g 1034 t
ORIGIN

Alignment Scores:
Pred. No.: 3,04e-252 Length: 4980
Score: 8285.00 Matches: 1590
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-617-099B-1 (1-1590) x E63816 (1-4980)

QY	1	MetSerAlaProLeuGlyProArgGlyProAlaProAlaProAlaSerGlnPro	20
Db	32	ATGTCGGCTCCGCTCGGGCCCGGGCGCCGCCCTCCACCCCGGGCGCTCTCAACCT	91
QY	21	ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysIleIle	40
Db	92	CCTCCGAGCCCGGATGTCGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	151
QY	41	LeuAlaValMetAspArgGlnLysGluGluGluGluGluGluGluGluGluGluGlu	60
Db	152	CTGGCTGTCATGGATCGTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	211
QY	61	LysGluGluHisLysAlaGlnProThrGlnTyrPheProPheSerGlyIleThrGluLeu	80
Db	212	AAAGAGAAACAAAGACCAACCCGACACAGTGGTTCCTTTAGTGGGATCACTGAACTG	271
QY	81	ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys	100
Db	272	GTAATAACGTTCTGCAGCCCCCAGCAAAACACCCATGAGAGAGAGAGAGAGAGAG	331
QY	101	LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln	120
Db	332	CTGCACCAACAATTTGAAATGTAAGAGAGCAAGTCAAGAGAGATGGGAGAGAGATCGCAG	391
QY	121	GlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe	140
Db	392	CAGCAGCAAGAGCAGAGAGGGTGTGTCGCCGACCTGTGGCATCTGCCACAGACAAAATTT	451
QY	141	AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly	160
Db	452	CGAGATGATCGGCCATATTTCTTATTTGCCAAACCAAGATTCGTGCTCGATGGGA	511
QY	161	GlyArgValSerLeuArgSerAsnLysValMetTyrValCysAsnLeuCysArgLysGln	180
Db	512	GGTCGAGTGCTTTACGCTCAACAGAGTTATGTTGGGTGTGTAATTTGTCGCGGAACAA	571
QY	181	GlnGlnLeuLeuThrLysSerGlyAlaTyrPheTyrAsnSerGlySerAsnThrLeuGln	200
Db	572	CAAGAAATCTCTCACTAAATCAGGAGCATGGTTTATATAATAGTGGGTCTAACACACTGCAG	631
QY	201	GlnProAspGlnLysValProArgGlyLeuArgAsnGluAlaProGlnGluLysLys	220

Db 632 CAACCTGATCAAAAGGTTCTCGAGGGCTTCGAAATGAGGAAGCCCTCAGGAGAGAA 691
Qy 221 AlalysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAla 240
Db 692 GCAAACTACAGCAGCAGCCAGTTCCAGAGAGCCCGAGTGACTTATCAGTACCTGCA 751
Qy 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
Db 752 GTTGAGAAAGCCGAGCTCATGGGCTCACAAGACAGGATACTATTAAAAATGGATCAGGA 811
Qy 261 VallysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
Db 812 GTGAAGCACCAGATGGCCAGTGACATGCCCTTCAGACAGAAAAACGAAGTCCATCAGTGTC 871
Qy 281 ArgAspGlnAanArgArgTyrGluGlnSerGluArgGluAspTyrSerGlnTyrVal 300
Db 872 AGGATCAAAATCGAAGTACAGCAAGTGAAGAGAGAGACTCTCAGATATGT 931
Qy 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 320
Db 932 CCTTCAGATGTACAATGCCAAGATCTCTTCGGATTATGCTGATAGACGATCTCAGCGT 991
Qy 321 GluProGlnPheTyrGluGlnProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340
Db 992 GAGCCTCAATTTTATGAAGAACCTGGTCAATTTAAATTAACGGGATTTCTAAACAGGAGAGGC 1051
Qy 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgArgAspGluTyr 360
Db 1052 CATGACATTCCAAGAGATATATGTGGATGATGAGATGTGGAGAGCAGATGAATAT 1111
Qy 361 GluArgGlnArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
Db 1112 GAAAGACAAAGCAGAGAGAGGAATACCAGGACGCTACAGAGTGTATCCAAATCTGGCC 1171
Qy 381 ArgTyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValSer 400
Db 1172 CGGTATCCGTAAAGCCCAACCCCTACGAAGACAAATGCGCATCCACGCTGAGGTGTC 1231
Qy 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuAsp 420
Db 1232 AGGACACGACATGAGAGAGGCACAGTATGTTCTTTGGCAACGCTGAACTAGAAGAT 1291
Qy 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
Db 1292 TCCAGGATTTCTCTGTAAGATGATGATAGACCATCAGGCAAGATCTGTTATCTGAAGT 1351
Qy 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
Db 1352 AGAGCTGCATGGAATGGAACCAACGATCGTATTCATATGGAAGAACTCCAGAGAGCTCAGGGA 1411
Qy 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
Db 1412 CAAAGTCTTATCACAAGAGACCTCAAAATCATAGTCTCCACCCCTCGCGGAGGCCCT 1471
Qy 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
Db 1472 ATACCGCTTGATACACAGACATGAGCGCGCTGACTCCCTACGGAACACACCACTTA 1531
Qy 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
Db 1532 GATCCAGCTCTGCTGTGAGGAAACCAAGCGAGGAGAAAAATGGAACCCATGTTAAGGAT 1591
Qy 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
Db 1592 GATCTTTGAGTTCAGACCATGCGAGTCAGTGAGCGGCCGCCACCAAGGCCCTCATAAA 1651
Qy 541 SerLysLysGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAla 560
Db 1652 TCCAGAAAGAGAGTAAATGCGCCAGGTTTCACTGACGACTCGAGAGAGAGCTGGCA 1711
Qy 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
Db 1712 TCCACACCTGATATACAGCTGTGATGATGTGGAGCTGGAAAGCGAGGTGTGAGTGAG 1771

Qy 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
Db 1772 AAAGGGACAGTCAAAAGGGHAAAAGAACTAGTGAGCAGGAGTTTTGTGCGATTCT 1831
Qy 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
Db 1832 AACACCAAGTCTGAGACACAAAGAAAAGGATGTACTATGTGTGCCACTCTTTGGAGAG 1891
Qy 621 AspLeuGlnTrpSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640
Db 1892 GATTTGGAATGGTCTGAGCCTCAGATTAAGGACTCTCGGGGTAGATACCTGTAGTAGACA 1951
Qy 641 ThrLeuAsnGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
Db 1952 ACCCTTAAAGAGGAGCATAGCCATAGTAGTAAGCACCTGTGACCTGGCAGCCATCCAAA 2011
Qy 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 2012 GATGAGATCGCTAAATTTGGTCTGATTTTATTAATAAGCGTTTAAAGATGGAGTGTA 2071
Qy 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
Db 2072 CCTCGAGACTCAGGAGCAATGCTGGGCTTAAAGGTTCTAGGAGCAAGATGACTGAATCA 2131
Qy 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 2132 GGTGCACTTTGTGCAATTTTATCCAAAGTAAAAAAGAAAGTTTAGTGTACTGTAGGA 2191
Qy 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr 740
Db 2192 CATCTTAGACAGGTGATGAAGTCTTGGATGGAATGGAGGCTATTGCAAGGAGCCACA 2251
Qy 741 PheGluGluValTyrAsnIleLeuLeuGluSerLysProGlnProGlnValGluLeuVal 760
Db 2252 TTTGAGGAGTTTACAACATTTCTAGAAATCCAAACCTCAACCAAGTTGAGCTTGT 2311
Qy 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
Db 2312 GTTTCAGGCGCAATGGAGATATCTTAGAATATCTGATAGCAGCATGCAACTGGAA 2371
Qy 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2372 TCCAGTTCTAGCTCATTTGAATCTCAAAAATGGACCGTCTCTCTATATCCGTTACTCA 2431
Qy 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2432 CCCATGAGTCTGTCATGCTGAGGAGTGTCCCGCAGTTCTTATCTGACAGCTTCAATA 2491
Qy 821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 2492 AAATATGGTTTGCAAGGTTGGTCACAGTTGATAGTTACAATTTTGGGAGCAAGGAT 2551
Qy 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
Db 2552 CTCCCTTCAGGAGAGATGGAGGCCAGGAGATCTTATGTATAGATTACTTCTTCCCA 2611
Qy 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
Db 2612 GATAGAGTGTATAAAAATAAGAGAGAAACAAAACACAGTCAAGAAAACCTTTGGAAACCCAAA 2671
Qy 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2672 TGGAACACAGACTTTCATTTTATCTCTGTCACCGAAGAGATTCCTGTGACAGATCTG 2731
Qy 901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu 920
Db 2732 GAATATACCTTTTGGGATCAAGCTAGAGTTCGAGAGAGAGAGAGCAATTTCTTAGGAG 2791
Qy 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTyrLysLeuGln 940
Db 2792 ATTTTAATTCGAATGGAAACAGCTTTGCTAGATGATGAGCGGCACCTGGTATATAGCTGAG 2851

941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgGln 960
2852 ACCCATGATGTCCTCCATTCGACATCCCTCGCCCTTCCCATATATGCCCCGGAGGCAG 2911
961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
2912 CTCATGGAGAGCCCAACGCCAGCTGCCAAGGTCCAGAGATATAGTGCAGTGAA 2971
981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
2972 GTGTCTGACTACGACTGCGAGGATGCGTGGGAGTAGTGTGAGATTATGCACAAATGC 3031
1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerHisCys 1020
3032 CGCGATCTTCAAAAGCTCCAGCTTGTGCGTCCGACAAAGATCATGTGCATCAATCATTC 3091
1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
3092 TCACCATCAGGGTCTCTCATCGAGTAGATGTTATAGGAGGACAGGTCAATGTCCTCT 3151
1041 SerAlaProProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
3152 AGTGCCCTCTCTCTCAAGGAATGTGGAACAGGGCCACCGAGGGACAGTGTACTGGC 3211
1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
3212 CATTAACAACACATTAAGCGAATGGATAGACACCGTGTATGGATGACCACTACTCTTCA 3271
1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
3272 GATAGAGACGGGATTGTGAGCAGCAGATAGACAGCCATATCACAGATCCAGATCACA 3331
1101 GluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProAsp 1120
3332 GAACAACGGCTCTCTAGAGCGACCAACCCCGCTCCAGATCTCTGAAGCTCTGAT 3391
1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140
3392 ACAAACTCATGAGGTGATGCTCTTCAATGACTGGAGATCTGCCCTCTCTCACCT 3451
1141 AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrPro 1160
3452 GCCTTATCAGAGGTCTCACCTCTGACCGGTCTGTCCAGACAGCCCATCAAGTACTCCG 3511
1161 GlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArg 1180
3512 GGAACAGAGCAGAGGGCCGACAGCTTCCACAGCTTCCACAAAGGACATTTGGAGAGA 3571
1181 SerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnVal 1200
3572 AGTGCTATGATATAGAGAGAGAAATCGCCAAATGAAATCTTAACAAATACAAACAGGTA 3631
1201 AlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpAsp 1220
3632 CGCGATTCAGACCCAGACTGGAGCAGATTAACATTCGAAGTATCGCTCAGATGGGAT 3691
1221 ProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSerAspVal 1240
3692 CCACATAGAGGGGCAGTACTGTTCCACTAATCTCCGACAGTGTAGTAACTGATGTA 3751
1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerVal 1260
3752 TCTCGCGTTTCAAGGACTAGTAGTCTTCTCGTTTCAGCAGCAGCAAGTACATGTCCTGC 3811
1261 GlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsn 1280
3812 CAATCAGAGCGCGCAGAGAGAAACAGGAAATCAGTGTCTTTTATCATCCAAATGCAAAAC 3871
1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMet 1300
3872 AGACAGATGGCGTGTGGGGAAGAACTTACCAAAAGCCAGCATCAGTGAGACATG 3931
1301 CysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThr 1320

3932 TGCTCATCTGGAGAGAAATGACGGCAGCCAGTCCGACATGTCAGTGGGGCCCTCTGGGTACC 3991
1321 SerGlyLysLysArgArgSerSerIleGlyValLysMetValAlaIleValGlyLeuSer 1340
3992 AGTGGCAAGAGCGCGCATCTAGCAATTGGGGCAAAATGGTAGCTATTGTTGGTCTCTCA 4051
1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArg 1360
4052 CGGAAAGTTCGAGTGCCTCTCAACTCAGCCAAACCCGAAGGAGGAGGTAAAGAGCTACGG 4111
1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
4112 AGCATCTGTTTCAGAGAAGCAGGAGACCGGGCTAGCAGTGGAGATGAGAACTGGATGACC 4171
1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsn 1400
4172 CGCAGGCGCAGCGCGGAATCCACAGATGGCAGCATGAACAGCTATAGCTCGGAGGAAT 4231
1401 LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly 1420
4232 CTGATCTTCCCTGGGGTCCGCTCTGACAGCCAGTTCAGTGAATTTCTCGATGTC 4291
1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440
4292 CTGGGCCCTGCTCAGCTAGTGGGACCGCAGACCCCTGGCTACTCTGCAATGGGTGACATT 4351
1441 GlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAlaArgGly 1460
4352 CAGTGGGGAATGATGATAAAAGGACAGCTGGAGGTAGAATCATCCGGGCGCGCGC 4411
1461 LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu 1480
4412 CTGTGTGTAACCAAGGTTCCAAAGACATGCCAGACCGTATGTCAAGGTATCTGTTA 4471
1481 AspAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeuGluPro 1500
4472 GACAAACGGAGTCTGATAGCCAAAGAAACCAAGGTGGCGAGAAAGCCCTGGAGCCC 4531
1501 LeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleIle 1520
4532 CTGTACACGACGCTCTGTCTTCCTTCAGAGAGAGCCCGGAGGGGTGTACAGATCAT 4591
1521 ValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu 1540
4592 GTCTGGGAGATTATGTTGTTGATGATCACAATCTTTATGGAGTGGCCCGACATCTC 4651
1541 LeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPheProProSer 1560
4652 TTAGATGAATCGGAATATCCAACTGTTGATGGATGGATTCCTTCCCTCTCTCTCC 4711
1561 SerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGlu 1580
4712 TCCTAGTAGATCCAACTCGGACCTCTGACAGAGAGGCTTCCCATCTCTCTGGA 4771
1581 SerSerThrGlyProSerTyrSerArgSer 1590
4772 AGTTCTACGGACCTTCTTACTCTCTGTTCA 4801

RESULT 3

AB021131
LOCUS Mus musculus RIM2 mRNA, complete cds. 4800 bp linear ROD 20-DBC-2000
DEFINITION AB021131
ACCESSION AB021131
VERSION AB021131.1 GI:11611474
KEYWORDS Rim2.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Ozaki, N., Shibasaki, T., Kashima, Y., Miki, T., Takahashi, K., Ueno, H., Sunaga, Y., Yano, H., Matsuura, Y., Iwanaga, T., Takai, Y. and Seino, S.

TITLE	CAMP-GEII is a direct target of CAMP in regulated exocytosis
JOURNAL	Nat. Cell Biol. 2 (11), 805-811 (2000)
MEDLINE	20512528
PUBMED	11056535
REFERENCE	2 (bases 1 to 4800)
AUTHORS	Ozaki, N. and Seino, S.
TITLE	Direct Submission
JOURNAL	Submitted (12-DEC-1998) Nobuaki Ozaki, Nagoya University School of Medicine, First Department of Internal Medicine; Nagoya, Nagoya 466-8550, Japan (E-mail:n-ozaki@med.nagoya-u.ac.jp, Tel:81-52-744-2142, Fax:81-52-744-2157)
COMMENT	Sequence updated (22-Jun-1999) Sequence updated (02-Feb-2000).
FEATURES	Location/Qualifiers

```

BASE COUNT      1453 a    1103 c    1246 g    998 t
ORIGIN          KUFFSSVDFVLPAPLIRKASQSSSESSISGFSTSKS"

Alignment Scores:
Pred. No.:       4.2e-241           Length:         4800
Score:           7930.00            Matches:        1529
Percent Similarity: 96.16%          Conservative:    0
Best Local Similarity: 96.16%       Mismatches:     1
Query Match:      95.72%            Indels:         60
DB:               10                Gaps:           1

US-09-617-099B-1 (1-1590) x AB021131 (1-4800)

QY      1 MetSerAlaProLeuGlyProArgGlyArgProAlaProAlaAAsrGlnPro 20
Db      32 ATGTGGCTCCGTCTGGGCCCGCGGGCGCCCCGGCTCCACCCGGCGGCCTCTCAACCT 91
QY      21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLyseIle 40

```

QY 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuAsp 420
DB 1232 AGGCACACATGAGAGAGGACAGTGTATTTCTTTGGCAACCGCTGAACCTAGAGAT 1291
QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
DB 1292 TCCAGGATTTCTCTGCTAAGGATGATAGACCATCAAGGCAAGATGCTATCTGAACGT 1351
QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
DB 1352 AGAGCTGCAATGGAATAACCAACGATCGTATTCAATGGAAGAACTCGAGAGGCTCAGGGA 1411
QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
DB 1412 CAAGTCTTATCCCAAGAGGACCTCAATCATAGTCTCTCCACCCCTCGGGGGAGCCCT 1471
QY 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
DB 1472 ATACCGCTTGATAGACCATGAGCGCGCTGACTCCCTACGGAACAGCACCACTTA 1531
QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
DB 1532 GATCCAGCTCTGCTGTGAGGAAAAACGAAGCGAGAAAAAATGGAACCATGTTAAGGAAT 1591
QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
DB 1592 GATTTTGGATTTGAGACAGTCCAGTCCAGTCAAGTGGCGCCGCCACCAAGGCGCTCATAAA 1651
QY 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAla 560
DB 1652 TCCAGAAGAGAGGATAAATGCGCCAGGTTCCTAGCAGCTCGGAGGAGGAGCTGCA 1711
QY 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
DB 1712 TCCACACCTCAGTATACAGCTGTGATGATGAGCTGGAAGCGCAGAGTGTGAGTGA 1771
QY 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
DB 1772 AAAGGGGACAGTCAAAAGGAAAAAAGAACTAGTGAGCAGGAGTTTGTGCGATCT 1831
QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
DB 1832 AACACAGGCTGTGAGAGCAAAAGAAAGGATGATCTATGTTGGCCACTTTTGGAGAG 1891
QY 621 AspLeuGluTyrSerGluProGlnLysAspSerGlyValAspThrCysSerSerThr 640
DB 1892 GATTTGGAATGCTCTGAGCTCAGATTAAGGACTCTGGGGTAGATACCTGTAGTAGACA 1951
QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
DB 1952 ACCCTTAAGGAGGATAGCCATAGTATAGACCCCTGTGACCTGGCAGCCATCCAAA 2011
QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
DB 2012 GATGGAGATCGCTAATTTGCTGATTTTATTAATAGCGTTTAAAGATGGAGTGA 2071
QY 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyLysMetThrGluSer 700
DB 2072 CCTCGAGACTCAGGAGCAATGCTGGCTTAAAGGTTGTAGGAGGAAAGATGACTGAATCA 2131
QY 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
DB 2132 GGTGCACTTTGTGCAATTTATACCAAGTAAAAAAGGAAGTTTAGCTGATATCTGAGGA 2191
QY 721 HisLeuArgProGlyAspGluValLeuGluTyrAsnGlyValArgLeuLeuGlnGlyAlaThr 740
DB 2192 CATCTTAGCCAGGTGATGAGTCTTGGAAATGGAATGGAGGCTATTGCAAGAGGCCACA 2251
QY 741 PheGluGluValTyrAsnIleLeuLeuGluSerLysProGluProGlnValGluLeuVal 760
DB 2252 TTTGAGGAAGTTTACAAACATTTCTTGAATCCAAACCTGAAACCAAGTTGAGCTTGT 2311

QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
DB 2312 GTTTCAGGCCAATTTGAGATATTTCTAGATATCTGATAGTACGATGACCACTCGAA 2371
QY 781 SerSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
DB 2372 TCCAGTTCTAGCTCATTTGAATCTCAAAAAATGACCGTCTCTTATATCGTTACCTCA 2431
QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
DB 2432 CCCATGAGTCTGCGATGCTGAGGAGTGTCCCGAGTTCTTATCTGACAGCTTTCATA 2491
QY 821 LysLeuTyrPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
DB 2492 AAACCTATGCTTTGACAGGTTGGTCAACAGTTGATAGTTACAAATTTGGGAGCAAGGAT 2551
QY 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
DB 2552 CTCCCTTCCAGGAGATGCGAGGCCAAGGAATCTTATGTTAAGATTTACTTCTCTTCCA 2611
QY 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysValThrLeuGluProLys 880
DB 2612 GATAGAGTGTATAAAATAAGAGAAACAAAAACAGTCAAGAAACCTTTTGGAAACCCAAA 2671
QY 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
DB 2672 TGGAAACAGACTTTTCATTTATTTCTCTGTCACCGAGAGAAATTCCTGTGACAGATGCTG 2731
QY 901 GluIleThrLeuTyrAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGlu 920
DB 2732 GAAATTAACCTTTGGGATCAAGCTAGAGTTGAGAGAGAGAGAGAGCAATTTCTAGGAGAG 2791
QY 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspGluProHisTyrTyrLysLeuGln 940
DB 2792 ATTTTAATTAATTTGGAACAGCTTGTCTAGATGATGAGCCGCACTGGTATTAAGTGCAG 2851
QY 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
DB 2852 ACCCATGATGTCCTCTCATTTGCCACTCCTCGCCCTTCCCCTATCTGCCCGGAGGAG 2911
QY 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
DB 2912 CTCCATGAGAGAGAGCCCAACGCGAGGCTGCAAGGTCGAAGAGATAAGTGACAGTGA 2971
QY 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
DB 2972 GTGTCTGACTACGACTCGAGGATGGCGAGTAGTGTGAGATTAATTCGACACAATGCG 3031
QY 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
DB 3032 CGCGATCTTCAAGGCTCAAGTGTGCGTGCCAGACAGTCAAGTCAATCAATCATTTGC 3091
QY 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
DB 3092 TCACCATCAGGCTCTCTCTCATCGATAGTGTATAGGAGACCAAGGTCATGTCGCT 3151
QY 1041 SerAlaProProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
DB 3152 AGTGGCCCTCTCTCTCAAGGAATGTGGAACAGGGGACCCGAGGAGACGCTGCTACTGCG 3211
QY 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
DB 3212 CATTAACAACATTTAGCCGAATGGATAGACACCGTGTGATGATGACCATCTACTTCA 3271
QY 1081 AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
DB 3272 GATAGAGAC----- 3280
QY 1101 GluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProAsp 1120
DB 3280 ----- 3280
QY 1121 ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140

Db 3280 ----- 3280
 Qy 1141 AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrPro 1160
 Db 3281 -----AGGTCTTCACCTCGTACCGGGTCTGTCCAGACAAAGCCATCAAGTACTCCG 3331
 Qy 1161 GlyThrGlyArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArg 1180
 Db 3332 GAAACAGACGAGGGGGCCGACAGCTTCACAGCTTCACCAAGAGGAAACATTTGAGAGA 3391
 Qy 1181 SerAlaMetAspIleGluArgAsnArgGlnMetLysLeuAsnLysTyrlsGlnVal 1200
 Db 3392 AGTGCTATGATATAGAGAGAGAAATCCCAATGAACTTAAACAATACAAACAGTA 3451
 Qy 1201 AlaGlySerAspProArgLeuGlnAspTyrlHisSerLysTyrlArgSerGlyTrpAsp 1220
 Db 3452 GCCCGATCAGACCCAGACGTGAGCAAGATACCATTCGAAAGTTCGCTCAGGATGGAT 3511
 Qy 1221 ProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSerAspVal 1240
 Db 3512 CCACATAGGGGGCAGATGCTTTCCCACTAAATCCTCGACAGTGTATGATGATGA 3571
 Qy 1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerThrMetSerVal 1260
 Db 3572 TCTCGGGTTTCAAGACTAGTAGTCTCTCTGTTTCAGACAGACAGCTACATGTCCTC 3631
 Qy 1261 GlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsn 1280
 Db 3632 CAATCAGACGCGCCGAGAGAAACAGGAAATCAGTGTCTTTACATCCAAATGCAAAAC 3691
 Qy 1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerLysSerGlyAspMet 1300
 Db 3692 AGACAGATGGCGGTCTCGGGGAAGAACTTGACCAAAAGCACAGCATCAGTGGAGACATG 3751
 Qy 1301 CysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThr 1320
 Db 3752 TGCTCAGTGGAGAAATGACGGCAGCAGCTCCGACATGAGTGGGGCCCTCGGTACC 3811
 Qy 1321 SerGlyLysLysArgArgSerSertleGlyAlaLysMetValAlaIleValGlyLeuSer 1340
 Db 3812 AGTGCAAGAGCGCGCATCTAGCATTTGGGGCAAAATGGTAGCTATTGTTGGTCTCTCA 3871
 Qy 1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArg 1360
 Db 3872 CGGAAAGTCGAGTGCCTCTCACTACGCAAAACCGAAGGAGGAGGTAAAGATACGG 3931
 Qy 1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
 Db 3932 AGCACTGTTTCAGAGAGACCGGAGACCGGGCTAGCAGTGGAGATGAGGACTGGATGACC 3991
 Qy 1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrlSerSerGluGlyAsn 1400
 Db 3992 CGCAGGCGCAGCGGGAATCCACAGATGCGACATGAACAGCTATAGCTCGAAGAAAT 4051
 Qy 1401 LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly 1420
 Db 4052 CTGATCTTCCCTGGGGTCCGCGTGGCTCTGACAGCCAGTTCAGTGAATTTCTCGATGGC 4111
 Qy 1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440
 Db 4112 CTGGCCCTCTCAGCTAGTGGAGCGCAGACCTCGGCTACTCTCTGCAATGGGTGACAT 4171
 Qy 1441 GlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAlaArgGly 1460
 Db 4172 CAGTGGGAATGATGATAAAAGGACAGCTGGAGGTAGAAATCATTCGCGGCGCGCGC 4231
 Qy 1461 LeuValValLysProGlySerLysThrLeuProAlaProTyrlValLysValTyrlLeuLeu 1480
 Db 4232 CTTGTGGTAAACACAGGTTTCAAGACACTGCCAGCACCTGATGTCAAGGTGTATCTGTA 4291
 Qy 1481 AspAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeuGluPro 1500

Db 4292 GACAACGGAGTCTGCATAGCCAAAAAGAAACCAAGGTGGCGAGAAAGCCCTGGAGCCC 4351
 Qy 1501 LeuTyrlGlnGlnLeuLeuSerPheGluGlnSerProGlnGlyArgValLeuGlnIleIle 1520
 Db 4352 CTGTACCAGCAGCTCTTGTCTTCGAGGAGAGCCGCCAGGGGAGGGTGTACAGATCATT 4411
 Qy 1521 ValTrpGlyAspTyrlGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu 1540
 Db 4412 GTCTGGGAGATATATGGTCGTATGATGATCAAAATCCCTTTATGGAGTGGCCACATACTC 4471
 Qy 1541 LeuAspGluLeuGluLeuSerAsnMetValIleGlyTrpPheLysLeuPheProProSer 1560
 Db 4472 TTAGATGAACCTGGACATATCCAAATGATGATGATGATGATGATGATGATGATGATGAT 4531
 Qy 1561 SerLeuValAspProThrSerAlaProLeuThrArgAlaSerGlnSerSerLeuGlu 1580
 Db 4532 TCCCTAGTAGATCAACCTTGGCACCTCTGCAAGAGAGAGCTTCCCAATCGTCTCTGAA 4591
 Qy 1581 SerSerThrGlyProSerTyrlSerArgSer 1590
 Db 4592 AGTTCTACCGGACCTTCTTACTCTCGTTCA 4621
 RESULT 4
 AF199331
 LOCUS AF199331 5550 bp mRNA linear ROD 04-JUL-2000
 DEFINITION Rattus norvegicus RIM2-5B (Rim2-5B) mRNA, complete cds.
 ACCESSION AF199331
 VERSION AF199331.1 GI:8925875
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 5550)
 AUTHORS Wang, Y., Sugita, S. and Sudhof, T.C.
 TITLE The RIM/NIM family of neuronal C2 domain proteins. Interactions with Rab3 and a new class of Src homology 3 domain proteins
 JOURNAL J. Biol. Chem. 275 (26), 20033-20044 (2000)
 MEDLINE 20347919
 PUBMED 10748113
 REFERENCE 2 (bases 1 to 5550)
 AUTHORS Wang, Y. and Sudhof, T.C.
 TITLE Direct Submission
 JOURNAL Submitted (27-OCT-1999) Center for Basic Neuroscience, The University of Texas Southwestern Medical Center, 6000 Harry Hines Blvd., Dallas, TX 75235-9111, USA
 FEATURES
 Location/Qualifiers
 1..5550
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 1..5550
 /gene="Rim2-5B"
 52..4629
 /gene="Rim2-5B"
 /codon_start=1
 /product="RIM2-5B"
 /protein_id="AAF81653.1"
 /db_xref="GI:8925875"
 /translation="MSAPLGRGRPATPAASPPPPPEPDLHLTEERKTIQAVM
 DRQKSEKSGSVLKLHQFMYTKQVKGWSESSQOQBGDAFTCGIHKTKFPAD
 GCHNCSYCKTFCARCGVLSNKMVMVNLCTQOILTKSGAFVYNSGNTPO
 QPQKALRLSPSEAPQKAKLHQFQPGDSSVPAVERAGLTKRQDSIKNG
 SKMKHQAISDPSDRKSPSVSDRNDREYDQSEEREYSQVPSDSTWPSPSYADR
 RSQEPQPEPHLYRDNRRCHRHSHKEYIVDDDEVSDRYERORREHYQART
 SPMARIPTVQPIEYEMRIHAIVSPARHRSRHSVLSNAELSDRIILLNDRFS
 RQSVSRFAAMENQESYMERTEAQSSYPQRTNHSPTTPRPSPIPLDRPLER
 ADSLKGHLLDPSAVKTKRMETMLRNDLSLSDQSSVRRPPPHKSKGKMKR
 QVSLSSSEELASTPEYTCDDVIESEVGEKGMDEYSMLHASMHSSASPSLHP
 VTMQPSKDGRLIGRIILLKRLKDGSPVPSGMLGLKVVGKMTESGRICAFITKVK
 KGLADTVGHLRPGDEVLWNGRLQGAFTFEVYNIILSKRPQVLSVSRPQMDPV
 RISPSTHAQLESSESSSSSESQKMDRPSISVTPSPMSPGMLRDVQPLSQSLSTKLWFDKY

GHOLIIVTLGAKDLPSRGDRPNPYVKIYFLPDRSDKDKRRKTKVKKTLLEPKWNTFF
 IYSPVRRERERLEITLWDQARVREERSEFLEITLLEETALLDDEPHWKLQTHD
 VSSLLPLPHSPYMPRQLGCESTRRLQSKRISDSSEVDYDCEDGVGVSDYRHRGR
 DLQSTLSVEQVMSNHDSPGPHRVDVIGTRSRWSVPPPPQRNVSGLGKGTAT
 GHYNTISRMDRHLEMTDHYSSERDSHFLTLPRSHRQTSEHHRDGRDCEAADRQPYH
 RSTSTQRPFLERTTRSRSSERADTNLMSFPLMTGRSAPPSPALSRHPRTGSVQ
 TSPSTPTVTRRGKQLPQPPKQTLERSAMDIERNKQMLNKYKOVAGSDPRLQDY
 HSKYRSGMDPHRGADVTSTKSDSDVSAVSTSSASRFSSTSYMSVQSRPRGNR
 KISVFTSKWQSRQWYSGKSMKSTSLSGDMCSLEKXGDSQSTAYAGAGTSGKKRS
 SIGAKMVALVGLSRKRSASQLSQTEGGGKKLSTVQRSTETGLAVENRNMTRQASR
 ESTDGKNSYSSEGNLIFPGVRLASDSQSFDLGGLGPAQLVGRQTLATPAMGDIQV
 MMDKQGLYELTRALGLVVKPGSKTLPAPYVKVYLLDNGVCIATKKTIVARKTLEPL
 YQQLLSFSPQKQLQIIVMGDYGRMDKSFVNGVAQIILLDELELSNMVWGNPKLFPF
 SSLVDTPLATLRRASQSSLESSTGPSYSRS"

BASE COUNT 1663 a 1281 c 1388 g 1218 t
 ORIGIN

Alignment Scores:

Pred. No.: 2,93e-226 Length: 5550
 Score: 7459.50 Matches: 1452
 Percent Similarity: 91.38% Conservative: 21
 Best Local Similarity: 90.07% Mismatches: 30
 Query Match: 90.04% Indels: 109
 DB: 10 Gaps: 4

US-09-617-099b-1 (1-1590) x AF199331 (1-5550)

QY 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
 DB 52 ATGTCGGCTCCACTCGCGGCCCGCGGCCGCGCTCCACCCCGCGCGCTCTCAGCCT 111
 QY 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysIlelle 40
 DB 112 CCTCCGAGCCCGAGATGCCGACCTCAGCCACTCAGCAAGAGAGAGAGAGAGAAATCATC 171
 QY 41 LeuAlaValMetAspArgGlnLysLysGluGluGluLysGluGlnSerValLeuLysIle 60
 DB 172 CAGGCTGTATGATCGTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228
 QY 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
 DB 228 ----- 228
 QY 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
 DB 229 -----AAG 231
 QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120
 DB 232 CTGCATCAACAAATTTGAATGTATAGGAGCAAGTCAAGAAATGGAGAGAGAGAGAGAGAG 291
 QY 121 GlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
 DB 292 CAGCAGCAAG 351
 QY 141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
 DB 352 GCAGATGATGCGCGCATAACTGTTCATATTGCCAACCAAGTCTGTGCTCGTGTGGA 411
 QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
 DB 412 GGTGAGTGTTCATTACGCTCAACCAAGTTATGTGGGTGTGTAAATTTGTGCCGAGAACCA 471
 QY 181 GlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
 DB 472 CAAGAAATCTCTACTAAGTCGGCGGTGTGTTTACAATAGTGGGTCTAATACACCGCAG 531
 QY 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluCluAlaProGlnGluLysLys 220
 DB 532 CAGCCTGATCAAAAGGCTCTTCAGGAGCTTCGAGTGGAGAGAGAGAGAGAGAGAGAGAG 591
 QY 221 AlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAla 240
 DB ----- 1686

DB 592 GCAAAACTGCATGAGCAGACGAGTTCAGGAGACCCCGGTGACTCATCAGTACCTGCA 651
 QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
 DB 652 GTTGAGAGAGGTGAGCTCATGGCTCACAAGACAGATTTCTATTAAAGATGGCTCAGGA 711
 QY 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
 DB 712 ATGAAGCACAGATCCCATGCTGACATGCTTTCAGACAGAAAAGAGTCCGTGAGTGTCC 771
 QY 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluLysGluLysGluLysGlnTyrVal 300
 DB 772 AGGGATCAAAATCGAAGATACGACCAACAACTGAAGAAAGAGAGAGATATTTCACAGTATGT 831
 QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSerGlnArg 320
 DB 832 CCTTCAGATAGCACATGCTAGATCTCCATCAGATTTATGCTCATAGACGATCTCAACGT 891
 QY 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340
 DB 892 GAACTCAATTTTATGAGAACCTGATCATTTTAAATTATAGGATTTCTAACAGGAGAGC 951
 QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
 DB 952 CATAGACATTTCCAAAGAGTATTATTGACACGACGAGATGTGAGAGCAGAGATGAATAT 1011
 QY 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
 DB 1012 GAAAGACAAAGAGAGAGAGAGAGTACCAAGGACGCTACAGAAAGTCAATCAAAATTTGGCC 1071
 QY 381 ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer 400
 DB 1072 CGGTATCCGTTAAGCCACACCTTATGAAGAGCAAAATCGGATCCACCTCAAGTGTCC 1131
 QY 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
 DB 1132 CGGCGACGACACAGAGAGAGGACAGTGTATTTCTTGGCAATGCTGACTAGAGAT 1191
 QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
 DB 1192 TCCAGGATTTCTCTGCTAGGATGATAGACCATCAAGGCAAAAGATCTGTATCTGAACGT 1251
 QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
 DB 1252 AGAGCTGCAATGGAACACGAGATGCTATTCATGGAAGAACTCGAGAGCTCAGGA 1311
 QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
 DB 1312 CAAAGTTCTTATCCCAAGGACCAAAATCATAGTCTCTCTACCCCTCGCAGAGACCT 1371
 QY 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
 DB 1372 ATACCGCTCGATAGACGAGCTGAGCGGTGCGGCTCCCTACGGAACACACACACTTA 1431
 QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgan 520
 DB 1432 GATCCAGCTCTGCTGTAAAGGAAAACGAGGAGAGAAAATAATGGAACAAATGTTGGAAT 1491
 QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProProProHisLys 540
 DB 1492 GATTTCTTTGAGTTTCAGACGATCTGAGTGTAGGCGCCGCCACCAAGGCTCATAAA 1551
 QY 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAla 560
 DB 1552 TCCAGAAAGAGAGTAAATGCGCAGGTTTCTCAGCAGCAGCTCAGAGAGAGAGTGGCA 1611
 QY 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerSerGluSerValSerGlu 580
 DB 1612 TCCAGCCTGATATACAGCTGTGTGACGTGGAGATTGAAGCGAGAGCGTAGTGAG 1671
 QY 581 LysGlyAspSerGlnLysLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
 DB 1672 AAGAGAGACATGAG----- 1686

QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyHisSerLeuGluGlu 620
Db 1687 -----TACAGCTGGTGGAG 1701
QY 621 AspLeuGluTrpSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640
Db 1702 CATGGCTCTGGCAT----- 1716
QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
Db 1717 -----AGCAGTGAGCATCCCAAGTCTTTGCACCTGTGACCTGGCAGCCATCCAAA 1770
QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 1771 GATGAGATCGCTAATTTGGTGGTATTTTAAATAAGCGTTTAAAGATGGAGTGT 1830
QY 681 ProArgAspSerGlyAlaMetLeuGlyLysValValGlyGlyLysMetThrGluSer 700
Db 1831 CTTCCAGATTTCCAGGAGCAATGCTGGCTTAAAGGTTGTAGGAGGAAAGATGACTGAATCA 1890
QY 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 1891 GGTGACATTTGTGCAATTTATACCAAGTGAAGGAAAGGAGTTTACGTACTGTAGGA 1950
QY 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr 740
Db 1951 CATCTTAGACAGGTGATGAAGTCTTGGATGGATGGAGCTATTGCCAGGAGCCACA 2010
QY 741 PheGluGluValTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuVal 760
Db 2011 TTTGAGGAAGTTTACAACATTTATCTAGAAATCCAGGCTTGAACCAAGTTGAGCTTGT 2070
QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
Db 2071 GTTTCAGGCCGATTTGGAGATAGTCTAGAAATACCTGATAGCACTCATGCAACATCGGAA 2130
QY 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2131 TCCAGTTCTAGCTCATTTGAAATCTCAAAAATGGATCGTCCGTCTATATCGTTACCTCT 2190
QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2191 CCCATGAGTCTCTGGCATGCTGAGGAGATGTCACAGTCTTATCTGGACAGCTTTCAATA 2250
QY 821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 2251 AAACATATGGTTTGACAGGTTGGTCACAGTTAATAGTTACAAATTTTGGGAGCAAGGAT 2310
QY 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
Db 2311 CTCCCTCCAGGAGATGGGAGCCAGGATCTTATGTTAAATTTACTTCTTCCA 2370
QY 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysLeuLeuProLys 880
Db 2371 GACAGAAGTGATAAAAAACAGAGAAGAACAAAAACAGTCAAGAAAACTTTGGAAACCCAAA 2430
QY 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2431 TGGAAACCAAGCTTTTCATTTATTTCTCTCTTCCCGAAGAGAAATTCGGGAAACGAATGCTC 2490
QY 901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGlu 920
Db 2491 GAATCACCTTTGGGATCAAGCTCGAGTTCAGAGGAGAAAGGAAATTTCTAGGAGAG 2550
QY 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGln 940
Db 2551 ATTTTAAATGAATGGAAACAGCTCTGTATGATGATGAACACACTGGTACAACTTCAG 2610
QY 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
Db 2611 ACCCATGATGTCCTCTCATTTGCCACTCCCTTCCATATATATGCCACGAGACAG 2670

QY 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
Db 2671 CTGCATGGAGAGAGAGCCACACGAGTTTGCAGAGTTCAGAGGATTAAGTGCAGCGAA 2730
QY 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
Db 2731 GTATCCGACTATGACTGCGAGAGCGCTGGAGTGGTGTGAGATTACCGACACCATGGC 2790
QY 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
Db 2791 COTGATCTTCAAGCTCCACATTTATCAGTCCAGAACCAAGTCATCTCATCAACCATTTGC 2850
QY 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
Db 2851 TCACCATCAGGCTCTCTCATCGAGTAGATTTATAGAGAGACTAGGTCTATGGTCGCT 2910
QY 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
Db 2911 AGTGTCCCTCTCTCTCAAGGAAATGTGAAACAGGGCTTCGAGGACACAGTGTACTTGC 2970
QY 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
Db 2971 CATTACAATACATTTAGCCGATGGACAGACACCGTGTCTATGGACGACCATTTACTCTTCA 3030
QY 1081 AspArgAsp----- 1083
Db 3031 GAGAGACAGCTATTCTCTACTCTCTCTCTCGACACAGGACACAGCATGAGCAT 3090
QY 1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg 1098
Db 3091 CACCACAGGATGGAAGGATTTGTAGACAGACAGATAGACAGCCATATCAGACATCCAGA 3150
QY 1099 SerThrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArg 1118
Db 3151 TCAACAGAACCAACGGCTCTCTTAGAGCGGACCAACCCGCTCCAGATCTCTGAAAGT 3210
QY 1119 ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro 1138
Db 3211 GCTGATACAAACCTCTATGAGTCTGATGCTTATTATGATGGAAGATCTGCCCCCTCT 3270
QY 1139 SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer 1158
Db 3271 TCACCTGCTTATCGAGTCTCTCACCTCGCATGGCTCTGTCCAGACCAAGTCCGTTCAAGT 3330
QY 1159 ThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeu 1178
Db 3331 ACTCCCTGTAAACAGACCGAAGCGGCGGAGCTTCTCTCAGCTCCCAACCAAGGAAACATG 3390
QY 1179 GluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLys 1198
Db 3391 GAAAGAGTGTCTATGGATATAGAGAGAGAAATCGCAAAATGAAACTTAAACAAATACAAA 3450
QY 1199 GlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysValTyrArgSerGly 1218
Db 3451 CAGGTAGCCGATCAGATCCAGACTGGAACAGATTAACATTCGAGTATCGTCTCAGGA 3510
QY 1219 TrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSer 1238
Db 3511 TGGGATCCACATAGAGGGGAGATCTGTTTCTACTAAATCTCTCGGACAGTGAATGAAT 3570
QY 1239 AspValSerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMet 1258
Db 3571 GATGTATCTCGGTTTCAAGGACTAGTAGTGTCTCTCTGTTTCAGCAGACCAAGTACATG 3630
QY 1259 SerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMet 1278
Db 3631 TCCGTCCATCAGAACCGCCGAGGAGAAACAGGAAATCAGTGTCTTTAGTCCAAATG 3690
QY 1279 GlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGly 1298
Db 3691 CAAAGCAGACAGATGGGTGTGTGGGGAGAGCATGCCAAAAGCACCAGCATCAGTGA 3750
QY 1299 AspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeu 1318

781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
2131 TCCAGTTCCTAGCTCAATTCCTCAAAATCGATCGTCCGCTATATCGGTACCTCT 2190
801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
2191 CCCATGAGTCTGGCATGCTGAGGGATGCCCCACAGTTCCTATCTGGAAGCTTTCAATA 2250
821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
2251 AAACATATGTTTGACAGGTGTGTACACAGTTAATAGTTACAAATTTTGGAGCCAAAGGAT 2310
841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
2311 CTCCTCTCCAGGAGATGGAGGCCCAAGGAATCCTTATGTAAATTTACTTCTCTCCA 2370
861 AspArgSerAspLysAsnLysArgThrLysThrValLysLysThrLeuGluProLys 880
2371 GACAGAAGTGATAAAACAGAGAGAAACAAACACAGTCAAGAAACTTTTGAACCCAAA 2430
881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
2431 TGGAACCCAGACTTTCATTTATTTCTCTGTTCCCGAGAGAAATTCGGGAAACGAATGCTC 2490
901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu 920
2491 GAAATCACCTCTTGGGATCAAGCTCGAGTTCGAGAGAGAAAGTGAATTCCTAGGAGAG 2550
921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpLysLysLeuGln 940
2551 ATTTTAAATGAAATGGAAACAGCTCTGTAGATGATGAACCCACACACGTTGACAACTTCAG 2610
941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgGln 960
2611 ACCCATGATGTCCTCTCATTTGCCACTCCCTCACCTTCTCCATATATGCCAGGAGACAG 2670
961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
2671 CTGCATGGAGAGAGCCCAACACGACGTTTGCAAAGTGCAGAGAGAAATAGTGACACGCAA 2730
981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
2731 GTATCCGACTATGACTGCCAGGAGCGCGTGGGAGTCGTCTCAGATTACCGACACGATGCG 2790
1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
2791 CGTGATCTTCAAAGCTCCACATTATCAGTGCCAGAACAAAGTCAATGTCATCAACCACTTC 2850
1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
2851 TCACCATCAGGCTCTCTCATCGAGTAGATGTTATAGGAAGGACTAGGTCTATGTCGCT 2910
1041 SerAlaSerProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
2911 AGTGTCCTCTCTCTCAAGGAATGTGGACACGGGCTTCGAGGGACACAGTGTCTACTGGC 2970
1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
2971 CATTACAAATCAATTAAGCCGATGGACAGACACCCGTGTATGGACCAACCACTACTCTTCA 3030
1081 AspArgAsp----- 1083
3031 GAGAGAGACAGTCATTTTCTCACTTACCTCGCTCTCGACACAGGCGAGCAGTGAGCAT 3090
1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg 1098
3091 CACCACAGGAGTGGAGGGATTTGGAAGCAGCAGATAGACAGCATATCACAGATCCAGA 3150
1099 SerThrGluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArg 1118
3151 TCACAGAACACAGGGCTCTCTCTAGAGCGGACCAACCCGCTCCAGATCCTCTGAACGT 3210

1119 ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro 1138
3211 GCTGATACAAACCTCATGAGTTCGATGCTTTCTTAATGACTGGAGATCTGCCCTCTCT 3270
1139 SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer 1158
3271 TCACCTGCTTATCGAGGTCTCACCTCGCACTGGCTCTGTCCAGACCAAGTCCGTCAGAT 3330
1159 ThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeu 1178
3331 ACTCCCGTAAACAGGACGAAGCGCGGCGAGCTTCTCTCAGCTCCCAACAAAGGACACATG 3390
1179 GluArg-----SerAlaMetAsp 1184
3391 GAAAGATGATCACAGAAAGACATGAGCTCCCAAGAAAGAGAACTCAGGTGCTATGGAT 3450
1185 IleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaGlySerAsp 1204
3451 ATAGAGGAGAGAAATCGCCAAATGAACTTAACTTAACAAATCAACAGGTAGCCGATCAGAT 3510
1205 ProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpAspProHisArgGly 1224
3511 CCCAGACTGGAAACAGATTAACATTCGAAGTATGCTCAGGATGGATCCACATAGAGGG 3570
1225 AlaAspThrValSerThrLysSerSerSerSerSerSerSerSerSerSerSerSerSer 1244
3571 CGAGATATCTGTTCTACTATAATTCCTCGGACAGTGAATGAATGATGATGATGATGATGAT 3630
1245 ArgThrSerSerAlaSerArgPheSerThrSerTyrMetSerValGlnSerGluArg 1264
3631 AGGACTAGTAGTCTCTCTGTTTCAGCAGCAGCAGTACATGCTCCATTCAGAACGG 3690
1265 ProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsnArgGlnMetGly 1284
3691 CGAGAGGAAACAGGAAATCAGTGTCTTTACGTCCAAATGCAAAAGCAGACAGATGGT 3750
1285 ValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMetCysSerSerLeuGlu 1304
3751 GTGTCGGGGAAGAGACGTGGCCAAAAGCAGCAGCATCAGTGGAGACATGTGCTCACTGGAG 3810
1305 LysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyLysLys 1324
3811 AAGAATGATGGCAGCAGTCGAGCAGCAGTGGGCGGCCCTGGCCACCAGTGGCAGAGAG 3870
1325 ArgArgSerSerIleGlyAlaLysMetValAlaIleValGlyLysSerArgLysSerArg 1344
3871 CGCGCATCTAGCATTTGGGCGCCAAATGCTAGTATTGTTGTTCTCTCAGGAAAGTCCG 3930
1345 SerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLysLeuArgSerThrValGln 1364
3931 AGTGCTCCCAACTCAGCCAAACCGAAGGAGGAGGTAAAGACTACGACGACCTGTCAG 3990
1365 ArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSer 1384
3991 AGAAGACAGGAGACCGGCTAGCAGTGGAGATGAGAACTGGATCAACAGACAGACGCGCAG 4050
1385 ArgGluSerThrAspGlySerMetAsnSerTyrSerSerSerGluGlyAsnLeuIlePhePro 1404
4051 CGGGAATCCACGATGGCAGCATGAACAGTTACAGTCAAGTCAAGAGGAAATCTGATCTTCT 4110
1405 GlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLeuGlyProAla 1424
4111 GGGGTCCGCTGGGCTCTCAGCAGCCAGTTCCAGCATTTCTTGAATGCTCTGGGCTCTGT 4170
1425 GlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMet 1444
4171 CAGCTAGTGGGAGCGCCAGACCTCTGGGACTCTCTGCAATGGGTGACATTCAGGTGGGAGATG 4230
1445 MetAspLysLysGlyGlnLeuGluValGluIleLeuArgAlaArgGlyLeuValLys 1464
4231 ATGGATTAAGAGGAGACAGCTGGAGTAGAGATCATCCGGGCTCGAGGCTCTGTAGTAAA 4290
1465 ProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeuAspAsnGlyVal 1484

QY 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGlnLysLys 220
DB 532 CAGCCTGATCAAAAGGCTCTTCAGGGCTTCGAGTGAGAAAGCCCTCAGGAGAGAG 591
QY 221 AlaLysLeuHisGlnGlnProGlnPheGlnGlnAlaProGlyAspLeuSerValProAla 240
DB 592 GCAAACTGCATGACGACGAGCTTCAGGGACCCCGGTGACTCATCAGTACCTGCA 651
QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
DB 652 GTTGAGAGAGTCCAGCTCATGGCTCAACAGACAGGATTCATTAAAGAAATGGCTCAGGA 711
QY 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
DB 712 ATGAGACACCATGCCAGTGCATGCTTCAGACAGAAAGATCCGTCAGTGCTC 771
QY 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrVal 300
DB 772 AGGGATCAAAATCGAAGATACGACCAAACTGAAGAAAGAGAGGAATATTACAGTATGTT 831
QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 320
DB 832 CCTTCAGATAGCACATGCTAGATCTCCATCAGATTATGCTCATAGACGATCTCAAGCT 891
QY 321 GlnProGlnPheTyrGlnLysProGlyHisLeuAsnTyrArgAspSerAsnArgGly 340
DB 892 GAACCTCAATTTATGAAGAACTGATCATTTAAATATAGGGATTCCTAACAGGAGAGGC 951
QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
DB 952 CATAGACATTCCAAAGAGTATTTGTAGACGCGAAGATGTGGAGACGACAGATGAATAT 1011
QY 361 GluArgGlnArgArgGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
DB 1012 GAAAGACAAAGAGAGAGGAGGTACAGGACGCTACAGAGTGCATCAAAATTTGGGC 1071
QY 381 ArgTyrProValLysProGlnProTyrGluGlnMetArgIleHisAlaGluValSer 400
DB 1072 CGGTATCCGGTAAAGCCACACCTATGAAGAGCAAAATCGGATCCACGCTGAAGTGTC 1131
QY 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
DB 1132 CGGGACGACACAGAGAGGACAGTGTATTTCTTGGCAATGCTGAATAGAT 1191
QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
DB 1192 TCCAGGATTTCTGCTAAGATGGATAGACCATCAAGCAAAAGATCTGTATCTGAAGCT 1251
QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
DB 1252 AGAGCTGCATGGAAACACAGCGATCGTATTCAATGGAAGAACTCGAAGCTCAGGGA 1311
QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
DB 1312 CAAAGTTCTTATCCACAAAGGACCACAATCATAGTCTCTCCATCCCTCGCAGGACCT 1371
QY 481 IleProLeuAspArgProAspMetArgAlaAspSerLeuArgLysGlnHisLeu 500
DB 1372 ATACCGCTCGATAGACGAGCTGAGGCGTGCAGCTCCCTACGGAACCAACACCACTTA 1431
QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
DB 1432 GATCCCGAGCTCTCTGTAAAGGAAACGAAAGCGAGAAATAATGGAACAATGTTGAGGAAT 1491
QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
DB 1492 GATTTTGGATTTCAGACCACTGTGAGTCACTGAGGCGCCGCCCCACCAAGGCTCTATAA 1551
QY 541 SerLysLysGlyLysMetArgGlnValSerLeuSerSerSerSerGluGluLeuAla 560
DB 1552 TCCAAGAAAGAGGATTAATCGCCAGGTTTCACTGAGCAGCTCAGAGGAGGAGTTGGCA 1611
QY 561 SerThrProGluTyrThrSerCysAspValGluLeuGluSerValSerGlu 580

DB 1612 TCCAGCCTGATATACAGCTGTGATGACCTGGAGATTGAACGAGAGCGTAGGTGAG 1671
QY 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlnValLeuSerAspSer 600
DB 1672 AAAGGAGACATGGAG- 1686
QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
DB 1687 -----TACAGCTGGTGGAG 1701
QY 621 AspLeuGluTyrPsrSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640
DB 1702 CATGGCTTGGCAT----- 1716
QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
DB 1717 -----AGCAGTGAGCATCCCAATGCTTTGCACCCCTGTGACCTGACCTGCAGCACCATCCAAA 1770
QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
DB 1771 GATGGAGATCCCTAAATTTGGTCTGATTATTAATAAGCCGTTTAAAGATGGGAGTGA 1830
QY 681 ProArgAspSerGlyValAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
DB 1831 CCTCGAGATTGAGGAGCAATGCTGGCTTAAAGTTGTAGAGGAAAGATGACTGAATCA 1890
QY 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
DB 1891 GGTGCAGTTTGTGCACTTTATTACCAGTGAAGGAGGAGGAGTTAGCTGACTGTAGCA 1950
QY 721 HisLeuArgProGlyAspGluValLeuGluTyrAsnGlyArgLeuLeuGlnGlnIleAlaThr 740
DB 1951 CATCTTAGACAGGATGATGAAGTCTGGAATGGAATGGAAGCTATTCCAGGAGGCCACA 2010
QY 741 PheGluGluValTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuVal 760
DB 2011 TTTGAGGAGTTTACAACTATTCTAGATCCAGGCTGAAACCAAGTTGAGTTGTT 2070
QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
DB 2071 GTTTCAGGCGGATTTGGAGATATGCTAGATATCTGATAGTACCTCATGCACCACTG 2130
QY 781 SerSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
DB 2131 TCCAGTTCTAGCTCATTTGAATCTCAAAAAATGGATCGTCCGCTCATATATCGGTTACTCT 2190
QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSer--- 819
DB 2191 CCCATGAGTCTGGCATGCTGAGGAGTGTCCACAGTTCTTATCTGGACAGCTTTCAGC 2250
QY 820 -----IleLysLeuTrpPhe 824
DB 2251 CAAAGCCTTAGTAGAAGAACCAAGCCTTTTGTCTCTAGGTTTCAGATAAAACTATGGTTT 2310
QY 825 AspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAspLeuProSerArg 844
DB 2311 GACAGGTTGGTTCACAGTAAATAGTTACAATTTTGGGAGCAAGAGTCTCCCTTCCAGG 2370
QY 845 GluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuProAspArgSerAsp 864
DB 2371 GAAGATGGAGGCCAAGGAAATCTTATGTTAAATTTACTTCTTCCAGACAGAGTGAAT 2430
QY 865 LysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLysTrpAsnGlnThr 884
DB 2431 AAAACAAGAGAGAAACAAAACAGTCAAGAAAACCTTTGGAAACCAATGGAACCACT 2490
QY 885 PheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeuGluIleThrLeu 904
DB 2491 TTTCAATTTCTCTCTGTTTCCAGGAGAGAAATTCGGGAGACCAATGCTCGAAATCACTT 2550
QY 905 TrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGluIleLeuIleGlu 924

Db 2551 TGGGATCAAGCTCGAGTTCGAGAGCAAGAAAGTGAATTCCTAGGAGAGATTTTAATTGAA 2610
Qy LeuGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGlnThrHisAspVal 944
Db 2611 TTGGAACAGCTCTGTAGATGATGAACACACACACACACACACACACACACACACACACAC 2670
Qy 945 SerSerLeuProLeuProAspProSerProTyrLeuProArgArgGlnLeuHisGlyGlu 964
Db 2671 TCCTCAATGGCACTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2730
Qy 965 SerProThrArgArgLeuGlnArgSerLysArgLysSerAspSerGluValSerAspTyr 984
Db 2731 AGCCCAACACGCAAGTTCGAAGGTGCAAGAGATAAGTGCACGCAAGTATCCGACTAT 2790
Qy 985 AspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGlyArgAspLeuGln 1004
Db 2791 GACTGCGAGACGCGTGGAGTCTGTGCAGTATACCGACACGATGGCGGTGATCTTCAA 2850
Qy 1005 SerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCysSerProSerGly 1024
Db 2851 AGCTCCACATATCAGTGCAGACACCAAGTCATGTCATCAAAACCATGCTCACCATCAGGC 2910
Qy 1025 SerProHisArgValAspValIleGlyArgThrArgSerTrpSerProSerAlaProPro 1044
Db 2911 TCTCTCATCGAGTAGATGTTATAGGAAGCACTAGGTCACTGCTCGCTAGTGTCCCTCCT 2970
Qy 1045 ProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGlyHisTyrAsnThr 1064
Db 2971 CCTCAAGAGATGTGGAAACAGGGCTTCGAGGGACACGTGCTACTGGCCATTACATACA 3030
Qy 1065 IleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSerAspArgAsp--- 1083
Db 3031 ATTAGCCGGATGGACACACACCGTGTCTATGGACGACCATTAATCTCTTCAGAGAGACAGT 3090
Qy 1083 ----- 1083
Db 3091 CATTTTCTCACTTACCTCGCTCTCGACACAGGCAGACAGTGCAGTACCAACAGGGAT 3150
Qy 1084 ---ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThrGluGln 1102
Db 3151 GGAAGGATGTTGTAAGCAGCAGATAGACAGCCATATCACAGATCCAGATCAACAGAAACA 3210
Qy 1103 ArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProAspThrAsn 1122
Db 3211 CGGCTCTCTAGAGCGGACACACCCGCTCCAGATCCTCTGAAGTGTGTATACAAAC 3270
Qy 1123 LeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerProAlaLeu 1142
Db 3271 CTCAATGAGTTCGATGCCCTTCATTAATGACTGGAAGATCTGCCCTCTTCACTGCCCTTA 3330
Qy 1143 SerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrProGlyThr 1162
Db 3331 TCGAGGTCTCACCCTCGCATGGCTCTGTCTGACAGCCAGTCCGTCAGTACTCCCGTAACA 3390
Qy 1163 GlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArg----- 1180
Db 3391 GGACGAAGGGCCCGGAGCTTCTCAGCTCCACCAAGGGAACCTGGAAAGATGATC 3450
Qy 1181 -----SerAlaMetAspIleGluGluArg 1188
Db 3451 ACAGAAGCATGGACTCCCAAGGAGAGAACTCAGGTGCTATGGATATAGAGAGAGA 3510
Qy 1189 AsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaGlySerAspProArgLeuGlu 1208
Db 3511 AATGCCAAATGAACATTAACAAATACAAACAGGTAGCCGATCAGATCCAGACTGGAA 3570
Qy 1209 GlnAspTyrHisSerLysTyrArgSerGlyTrpAspProHisArgGlyAlaAspThrVal 1228
Db 3571 CAAGATTACCAATTGCAAGTATCGCTCAGGATGGGATCCACATAGAGGGCGAGATCTGTT 3630
Qy 1229 SerThrLysSerSerAspSerValSerAspValSerAlaValSerArgThrSerSer 1248
Db 3631 TCTACTAAATCTCGGACAGTGTAAAGTGTATCTATCTGCGGTTTCAGAGCTAGTAGT 3690

Qy 1249 AlaSerArgPheSerSerThrSerTyrMetSerValGlnSerGluArgProArgGlyAsn 1268
Db 3691 GCTTCTCGTTTCAGCAGCACAAGCTACATGCTCGTCCATCAGACCGCCGAGAGAAAC 3750
Qy 1269 ArgLysIleSerValPheThrSerLysMetGlnAsnArgGlnMetGlyValSerGlyLys 1288
Db 3751 AGCAAAATCAGTCTCTTACGTCCTCAAAATGCAAAAGCAGACAGATGGTGTGTGGGGAAG 3810
Qy 1289 AsnLeuThrLysSerThrSerIleSerGlyAspMetCysSerLeuGlnLysAsnAspGly 1308
Db 3811 AGCATGCCCAAAAGCAGCAGCATCAGTGGAGACATGTGCTCTCCTGAGAAAGATATGATGCG 3870
Qy 1309 SerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGlyValLysArgArgSerSer 1328
Db 3871 AGCAGTCCGACACTGCACTGGGCGCTCTGGGACCCGTTGGCAAGAGCGCGCATCTAGC 3930
Qy 1329 IleGlyAlaLysMetValIleValGlyLeuSerArgLysSerArgSerAlaSerGln 1348
Db 3931 ATTGGGCGCAAAATGTTAGTATTGTTGTTCTCTCAGGAAAGTGCAGTGCCTCCAA 3990
Qy 1349 LeuSerGlnThrGluGlyGlyLysLysLeuArgSerThrValGlnArgSerThrGlu 1368
Db 3991 CTCAGCCAAACCGAAGGAGGAGTAAAAAGCTACGACGACTGTCCAGAGACGCGAG 4050
Qy 1369 ThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGlnAlaSerArgGluSerThr 1388
Db 4051 ACCGGCTAGCAGTGGAGATGAGAACTGGATGACAGACAGCCAGCCGCGAATCCACG 4110
Qy 1389 AspGlySerMetAsnSerTyrSerSerGluGlyAsnLeuIlePheProGlyValArgLeu 1408
Db 4111 GATGGCAGCATGAACAGTTACAGCTCAGAGGAAATCTGATCTTCTGGGGTCCGCTG 4170
Qy 1409 AlaSerAspSerGlnPheSerAspPheLeuAspGlyLeuGlyProAlaGlnLeuValGly 1428
Db 4171 GCCTCTCAGACCGAGTTTCAGGATTTCTTGATGCTGGGCTGCTCAGCTAGTGGGA 4230
Qy 1429 ArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnValGlyMetMetAspLysLys 1448
Db 4231 CGCCAGACCTCGCGACTCTCTGCAATGGGTGACATTCAGGTGGGATGATGATGAAGA 4290
Qy 1449 GlyGlnLeuGluValGluIleAlaArgGlyLeuValValLysProGlySerLys 1468
Db 4291 GGACAGCTGGAGGTAGAGATCACTCCGGCTCGAGGCTTGTAGTAAACCGAGTTCCAG 4350
Qy 1469 ThrLeuProAlaProTyrValLysValTyrLeuLeuAspAsnGlyValCysIleAlaLys 1488
Db 4351 ACATGCGCAGCACCATATGTCAAGGTGTACTCTGTAGACAAACGAGTCTCATAGCCAAA 4410
Qy 1489 LysLysThrLysValAlaArgLysThrLeuGluProLeuTyrGlnGlnLeuLeuSerPhe 1508
Db 4411 AGAAACCAAGGTGGGAGAGACGCTGGAGCCCTGTATCAGCAGCTGTTATCTTTTC 4470
Qy 1509 GluGluSerProGlnGlyArgValLeuGlnIleValTyrGlyAspTyrGlyValGlyMet 1528
Db 4471 GAAGAGACCCCGAGGAGGTGTACAGATCATTTGTCTGGGAGATNTAGCCGATG 4530
Qy 1529 AspHisLysSerPheMetGlyValAlaGlnIleLeuLeuAspGluLeuLeuSerAsn 1548
Db 4531 GATCACAAATCCCTTATGGGAGTGGCCAGATACTCTTAGATGAACTGGAATATCAAC 4590
Qy 1549 MetValIleGlyTrpPheLysLeuPheProProSerSerLeuValAspProThrSerAla 1568
Db 4591 ATGGTAATGGATGGTTCAAACTCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4650
Qy 1569 ProLeuThrArgArgAlaSerGlnSerSerLeuGluSerSerThrGlyProSerTyrSer 1588
Db 4651 CCTCTGACAGAAAGAGCTTCCCCAATCGTCTCTGGAAGTTCTACTGGACCTTCTTACTCT 4710
Qy 1589 ArgSer 1590
Db 4711 CGTTCA 4716

Db 952 CATGACATTCCAAAGAGTATATTTGTAGACGACGAAGATGTGGAGACGAGATGAATAT 1011
Qy GluArgGlnArgArgGluGluGlnGlnAlaArgTyrArgSerAspProAsnLeuAla 380
Db 1012 GAAGACAAAGAGAGAGAGAGATACCAAGCAGCCTACAGAAAGTATCCAAATTTGGCC 1071
Qy 381 ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer 400
Db 1072 CGGTATCCGGTAAAGCCACCAACCCCTATGAAGAGCAAAATGCGGATCCACGCTGAAGTGTCC 1131
Qy 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
Db 1132 CGGCACACACAGAGAGAGCAGTGTATTTCTTTGGCAAAATGCTGAACTAGAAGAT 1191
Qy 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
Db 1192 TCCAGGATTTCTCTGCTAGAGATGATAGACCAATCAGGCAAGATCTGTATCTGAAGCT 1251
Qy 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
Db 1252 AGAGCTGCAATGGAAACACAGCGATCGTATTCAATGGAAAGAACTCCGAGAGCTCAGGA 1311
Qy 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
Db 1312 CAAAGTTCTTATCCCAAAGAGCACCAATCATAGTCTCTTACCCCTCGCAGAGCCCT 1371
Qy 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisIleLeu 500
Db 1372 ATACCGCTCGATAGACAGAGCTGAGCGGTGCCACTCCCTACCGAAACAAACACACTTA 1431
Qy 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
Db 1432 GATCCAGCTCTGCTGTAAAGAAACCAAGCGAGAAATAATGGAAACAACTGTGAGGAAT 1491
Qy 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
Db 1492 GATTCTTTGAGTTCAGACCACTCTGAGTCAGTGGCGGCCCCACCACCAAGCCCTCATAAA 1551
Qy 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluGluLeuAla 560
Db 1552 TCCAAAGAAAGAGGTAAATGCCGAGTTTCATCTGAGCAGCTCAGAGAGAGGTGGCA 1611
Qy 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
Db 1612 TCCACGCTGAGTATACAGCTGTGATGACCTGGAGATTGAAGCGAGAGGTAGTGGAG 1671
Qy 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
Db 1672 AAAGGACACAGTCAAAAGGGGAAAGAAAGAACTAGTGAGCAGGCGAGTTTGGTCGACTCT 1731
Qy 601 AsnThrArgSerGluArgGlnLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
Db 1732 AACACCAAGCTGAGAGACAAAGAAAGATGATGATCTCTGTGGCCACTCTTTTGGACGAG 1791
Qy 621 AspLeuGluTyrSerGluProGlnLysAspSerGlyValAspThrCysSerSerThr 640
Db 1792 GATTGGATGGTCTGAGCCTCAGACTAGAGCTCTGSGGTAGACACCTGTAGTAGCACA 1851
Qy 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrThrGlnProSerLys 660
Db 1852 ACTCTTAACGAGGAGCGTAGCATAGTGATGAGCACCCTGTGACCTGGACCCATCCAAA 1911
Qy 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 1912 GATGAGATCCCTAAATTTGGTCTGATTTATTAATAAGCGTTTAAAGATGGGAGTGTA 1971
Qy 681 ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
Db 1972 CCTCGAGATTGAGAGCAATGCTGGGCTTAAGGTTTGAAGGAGGAAGATGACTGAATCA 2031
Qy 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 2032 GGTGCGATTTGTGCAATTTATTAACCAAGTGAAGAAAGGAAGTTTAGCTGACTGTAGGA 2091

Qy 721 HisLeuArgProGlyAspGluValLeuGluTyrAsnGlyArgLeuLeuGlnGlyAlaThr 740
Db 2092 CATCTTAGACCAAGGTGATGAGTCTTGGATGGATGGAGGCTATTGGCAGGAGGCCACA 2151
Qy 741 PheGluGluValTyrAsnIleLeuLeuGluSerLysProGluProGlnValGluLeuVal 760
Db 2152 TTTGAGGAAGTTTACACATTTATTAGATTCCAAAGCTGAAACCAAGTTGAGCTGTTT 2211
Qy 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
Db 2212 GTTTCAAGGCCGATTTGAGATATGCCTAGAATACCTGATAGCACTCATGCAACAACCTG 2271
Qy 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2272 TCCAGTTCTAGCTCATTTGATCTCAAAAATGGATGCTGCTCTATATGCTTACTCT 2331
Qy 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2332 CCCATGAGTCTGGCATGCTGAGGAGTGTCCACAGTTCTTATCTGCACAGCTTCAATA 2391
Qy 821 LysLeuTyrPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 2392 AAACATATGGTTTGACAGAGTTGGTCCACAGTTAATAGTTACAATTTTGGGAGCAAGAT 2451
Qy 841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
Db 2452 CTCCCTTCCAGGAGATGGAGGCCAAGGATCTTATGTTTAAATTTACTTCTTCCA 2511
Qy 861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
Db 2512 GACAGAGTGTAAAAACAAGAGAGAAACAACAACAGTCAAGAAACTTTTGGACCCAAA 2571
Qy 881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2572 TGGAAACAGACTTTTCAATTTATTTCTCTGTTTCCCGAAGAGATTTCCGGGACGAATGCTC 2631
Qy 901 GluIleThrLeuTyrAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu 920
Db 2632 GAATACACCTTTGGGATCAGCTCGAGTTCCGAGAGGAGAAAGTGAATTTCTTAGGAGAG 2691
Qy 921 IleLeuIleGluLeuGluThrAlaLeuLeuAspGluProHisIleTyrLysLeuGln 940
Db 2692 ATTTTAATTAATGGAAACAGCTCTCTTAGATGATGAACCAACACTGTGTACAACTCAG 2751
Qy 941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
Db 2752 ACCCATGATGCTCTCTCATTTGCCACTCCCTCACCTTCTCCATATATGCCACGAGACAG 2811
Qy 961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
Db 2812 CTGCATGGAGAGAGCCCAACACAGCGGTTGCAAGGTCGAGAGAAATAGGTGACAGCGAA 2871
Qy 981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
Db 2872 GTATCCGACTATGACTCGAGAGCGGCTGGAGTCTGTCTCAGATTACCAACACGATGGC 2931
Qy 1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
Db 2932 CGTGATCTTCAAGACTCCCATTTATGATGTCGAGCAAGTCAATGTCATCAACCAATTC 2991
Qy 1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
Db 2992 TCACCATCAGGCTCTCTCTCATCGAGTAGATGTTATAGGAGGACTAGGTCTATGTCGCT 3051
Qy 1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
Db 3052 AGTGTCCCTCTCTCTCAAGGAATGTGAAACAGGGCTTCAGGGACACGCTGTACTTGGC 3111
Qy 1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
Db 3112 CATTAACAATTAAGCCGATGGACAGACACCGTGTCTATGACGAGCACTTACTCTTCA 3171

QY 1081 AspArgAsp----- 1083
Db 3172 GAGAGAGAGCAGTCATTTTCTCACTTACCTCGCTCTCGACACAGCGACAGCAGTGGAGCAT 3231
QY 1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg 1098
Db 3232 CACCACAGGAGTGAAGGAGTGTGAAGCAGCAGATAGACAGCCATATCACAGATCCAGA 3291
QY 1099 SerThrGluGlnArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArg 1118
Db 3292 TCAACAGAACCAACGGCTCTCTTAGAGCGGACACACCCGCTCGAGATCCTCTGAAGCT 3351
QY 1119 ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro 1138
Db 3352 GCTGATACAACTCATGAGTTCGATGCTTCATTAATGACTGGAGATCTGCCCTCTCT 3411
QY 1139 SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer 1158
Db 3412 TCACCTGCCTTATCGAGGCTCTCACCTCGCACTGGCTGTGTCAGACAGTCCGTCAGT 3471
QY 1159 ThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeu 1178
Db 3472 ACTCCGTTACAGGACGAAGGGCGCGAGCTTCCTCAGCTCCCAACCAAGGGGAACACTG 3531
QY 1179 GluArgSerAlaMetAspIleGluArgAsnArgGlnMetLysLeuAsnLysTyrLys 1198
Db 3532 GAAAGA----- 3537
QY 1199 GlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGly 1218
Db 3537 ----- 3537
QY 1219 TrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSer 1238
Db 3537 ----- 3537
QY 1239 AspValSerAlaValSerArgThrSerSerAlaSerArgPheSerThrSerTyrMet 1258
Db 3537 ----- 3537
QY 1259 SerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMet 1278
Db 3537 ----- 3537
QY 1279 GlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGly 1298
Db 3537 ----- 3537
QY 1299 AspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeu 1318
Db 3537 ----- 3537
QY 1319 GlyThrSerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGly 1338
Db 3537 ----- 3537
QY 1339 LeuSerArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLys 1358
Db 3538 -----CGACGAGGTAAAAAG 3552
QY 1359 LeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrp 1378
Db 3553 CTACGAGCAGCTGTCACGAGAGCAGCAGGAGCGGCTAGCAGTGGAGATGAGGACTGG 3612
QY 1379 MetThrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGlu 1398
Db 3613 ATGACAAGACAGCCAGCCAGCGGGAATCCAGGATGGCAGCATGAACAGTTACAGCTCAGAA 3672
QY 1399 GlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeu 1418
Db 3673 GGAAATCTCATCTTCTCGGGTCCGCTCTGACAGCCAGTTCAGCGATTTCCCT 3732
QY 1419 AspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly 1438

Db 3733 GATGTCCTGGGCCCTGCTCAGCTAGTGGAGCGCCAGACCCCTGGGACTCTCTGCAATGGGT 3792
QY 1439 AspIleGlnValGlyMetMetAspLysGlyGlnLeuGluValGluIleAlaArgAla 1458
Db 3793 GACATTACGTTGGGGATGATGATGAAGAGGAGCAGCTGGAGGTAGAGATCATCCGGGCT 3852
QY 1459 ArgGlyLeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyr 1478
Db 3853 CGAGCCCTTGTAGTAAACACAGGTTCCAGACACTGCCAGCACCATATGTCAAGTGTAC 3912
QY 1479 LeuLeuAspAsnGlyValCysIleAlaLysLysLysThrLysValAlaAArgLysThrLeu 1498
Db 3913 CTGTTAGCAACCGAGTCTGATAGCCCAAAAGAAACCAAGTGGCGAGAGAGCGCTG 3972
QY 1499 GluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGln 1518
Db 3973 GAGCCCTTGTATCAGCAGCTGTATCTTTCAGAGAGAGCCCCAGGGGAGAGTGTACAG 4032
QY 1519 IleIleValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGln 1538
Db 4033 ATCATGCTGGGAGATTAATGGCGCATGATCACAATCTTTATGGAGTGGCCAG 4092
QY 1539 IleLeuLeuAspGluLeuGluLeuSerAsnMetValIleGlyTrpPheLysLeuPhePro 1558
Db 4093 ATACTCTTAGATGAATGGAATATCCAACTGATGTAATGGATGTTCAAACCTCTCCCT 4152
QY 1559 ProSerSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSer 1578
Db 4153 CCTCTCTCCCTAGTAGATCCAACTTGGCACTCTGACAGAGAGCTTCCCAATCGTCT 4212
QY 1579 LeuGluSerSerThrGlyProSerTyrSerArgSer 1590
Db 4213 CTGGAAGATTCTACTGGACCTTCTACTCTCTGTTCA 4248

RESULT 8
AF199329 4965 bp mRNA linear ROD 04-JUL-2000
LOCUS Rattus norvegicus RIM2-4C (Rim2-4C) mRNA, complete cds.
DEFINITION AF199329
ACCESSION AF199329
VERSION AF199329.1 GI:8925871
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 4965)
AUTHORS Wang, Y., Sugita, S. and Sudhof, T.C.
TITLE The RIM/NIM family of neuronal C2 domain proteins. Interactions with Rab3 and a new class of Src homology 3 domain proteins
JOURNAL J. Biol. Chem. 275 (26), 20033-20044 (2000)
MEDLINE 20347919
PUBMED 10748113
REFERENCE 2 (bases 1 to 4965)
AUTHORS Wang, Y. and Sudhof, T.C.
TITLE Direct Submission
JOURNAL Submitted (21-OCT-1999) Center for Basic Neuroscience, The University of Texas Southwestern Medical Center, 6000 Harry Hines Blvd., Dallas, TX 75235-9111, USA

FEATURES
source
1. 4965
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
gene
1. 4965
/gene="Rim2-4C"
CDS
52..4044
/gene="Rim2-4C"
/codon_start=1
/product="RIM2-4C"
/protein_id="AAF81651.1"
/db_xref="GI:8925872"

/translation="MSAPLGRGRPAPTPAASPPPPQPPMPDLSHLTEERKIIQAVM
 DRQKEEESQVLKHLQOPEMYKQVKKQGEESQOQKQGDAPTGGICHKTKFAD
 GCCHNSQCTQKCARCGRSVLSNKMVMVCLKQKQBEILTKSGAMFVNSGNTPO
 QPKALRGRLSEAPKAKLHEOTFOFGPPGDSVPAVERGRAGHLTRODSIKNG
 SGKHQIASMPDRKSPSVSDQMRVDSERREYSQVPSDSWMPSPSDYADR
 RSOREFOFTEPDLNTRNRGRHRSKSIYVDDSEVESRDEYERQREESTQARYR
 SDNLARYPAKQPIYEOEMRIHAEVSRARHERHSDVSLANAELESDRSISLMDRPS
 RQSVSERRAAMNORSYSMERTREAOQSSYPQRTTNHSPPTPRSPPLDPPELR
 ADSLRQHLDPSSAVKTRKEMETMLRNDLSLSDQSESVRPPPRPHSKKGGKMR
 QVLSSESEELAPTEVTSDDVEIESVGEKGDMEYSWLHSHSSSEASPMSLPH
 VYQPSKQDGLRILLNKLKDGSPVDRSGMLGLKVYGGKMTSREGLCAFTKVK
 KGLADTVGHLRDEGEVWNLQLOGATPEEVINILKSPKPPQVELVVRPIGMP
 RLPDTHAQLESSESSSESQSPSITSVPSMGLMDRVPQFLSQSLKLMFKFY
 GHQIVITLGAQKLPDSRGNRPYVYKIFLPSRDKNRRTKTKVKTLEPKWNTF
 IYGVPRRERFLEMLTTLMDQVRESEEFLEILLLELTALLDDEPHYKLOTHD
 VSSLPLPHSPYMRRLHGESPTRLQSRKISDSRVSDDCEDGVGVVYDRHGR
 DLQSSLTVPQVMSNNHSPSPSPRVDVIGTRSWSPSPVPPQNRVEQGLGRTAT
 GHNTISMDHRVMDHYSERDRDCEADROPYHRSRSTEQRLPLERTTTRRSSE
 RADTNLMRSPSLMGRSAPSPALSRHPRTGVSQVTSPTSPVTPURGRQLPOLPPK
 GTLERGGKKLRSTVQRSTETGLAVENRMNTQASRESTDGMNYSSEGNLIFFGV
 RLASDSQFSDFLGLGPAQLVGRQTLATPAMGDIQVGMMDKQGLVBIIRARGLVVK
 PGSKTLPAPVYKVVLLDNGVCIKKTKVARKTLEPLVYQQLSFPSPQSKVLIIVW
 GDYGRMDHKSFMVQAIIIDLELSNVMVGFKLPFPSSLVDTPLAPLTRASQSSLE
 SSTGPSYSRS"

BASE COUNT 1486 a 1145 c 1232 g 1102 t

ORIGIN

Alignment Scores:

Pred. No.: 1.67e-197 Length: 4965
 Score: 6543.00 Matches: 1283
 Percent Similarity: 81.82% Conservative: 18
 Best Local Similarity: 80.69% Mismatches: 29
 Query Match: 78.97% Indels: 260
 DB: 10 Gaps: 4

US-09-617-099B-1 (1-1590) x AF199329 (1-4965)

Qy 1 MetSerAlaProLeuGluGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
 Db 52 ATGTCGGCTCCACTCGGGCCCCGGGGCGCGCGGCTCCACCCCGGGCGCTCTCAGCCT 111
 Qy 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysLeile 40
 Db 112 CCTCCGAGCCGAGATGCCGACCTCAGCCACCTCAGGAAGAGAGAGAGAAATCATC 171
 Qy 41 LeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeuLysleile 60
 Db 172 CAGGCTGTCTATGGATCTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228
 Qy 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
 Db 228 ----- 228
 Qy 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
 Db 229 -----AAG 231
 Qy 101 LeuHisGlnGlnPheGluMetTyrlsGluGlnValLysLysMetGlyGluGluSerGln 120
 Db 232 CTGCATCAACAATTTGAATGTATAAGAGCAAGTCAAGAAATTTGGGAGAGAGATCAAA 291
 Qy 121 GlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
 Db 292 CAGCAGCAG 351
 Qy 141 AlaAspGlyCysGlyHisAsnCysSerTyrlsGlnThrLysPheCysAlaArgCysGly 160
 Db 352 GCAGATGGATCGGCCATTAAGTTCATATGCCAACCAGATTCCTGTGCTGTGTGGA 411
 Qy 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
 Db 412 GGTGAGTGTCTATACGCTCAACAAGAGTTATGTGGGTGTGTAATTTGTGCCGAGAAACAA 471

181 GlnGluIleLeuThrLysSerGlyAlaTrpPheTyrlsAsnSerGlySerAsnThrLeuGln 200
 Db 472 CAAGAAATCTCTACTAAGTCGGCGCGTGTGTTTACATAGTGGTCTTAATACACCGCAG 531
 Qy 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGlnLysLys 220
 Db 532 CAGCCTGTGATCAAAAGGCTCTTCGAGGCGCTTCGAGAGAGAGAGAGAGAGAGAGAG 591
 Qy 221 AlaLysLeuHisGlnGlnProGlnPheGlnGlyValAlaProGlyAspLeuSerValProAla 240
 Db 592 GCAAAATCTCATGAGCAGCAGCGAGTTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 651
 Qy 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrLysLysAsnGlySerGly 260
 Db 652 GTTGAGAGAGAGTCTCAGTCTCATGGCTCAACAAGCAGAGATTCATTAAAGATGCTCAGGA 711
 Qy 261 ValLysHisLeuIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
 Db 712 ATGAAGCAGCAGATCCCATGCTGATCATGCTCATGATGCTCATGATGCTCAACGT 891
 Qy 281 ArgAspGlnAsnArgArgTyrlsGlnSerGluGluArgGluAspTyrlsSerGlnTrpVal 300
 Db 772 AGGGATCAAAATCGAAGATACGACCAAGTGAAGAGAGAGAGAGAGATTCATGATGTT 831
 Qy 301 ProSerAspGlyThrMetProArgSerProSerAspTyrlsAlaAspArgSerGlnArg 320
 Db 832 CCTTCAGATAGCAATGCTGATGCTCATGATGCTCATGATGCTCATGATGCTCAACGT 891
 Qy 321 GluProGlnPheTyrlsGluGluProGlyHisLeuAsnTyrlsArgAspSerAsnArgArgLys 340
 Db 892 GAACTCTCAATTTATGAAGAACCTGATCATTTAAATATATAGGATTCCTAACAGAGAGGC 951
 Qy 341 HisArgHisSerLysGluTyrlsValAlaAspAspGluAspValGluSerArgAspGluTyrls 360
 Db 952 CATGACATTCCTCAAGAGATATATTGAGACGACCAAGATGTCGAGAGCAGAGATGAATAT 1011
 Qy 361 GluArgGlnArgArgGluGluGluTyrlsGlnAlaArgTyrlsArgSerAspProAsnLeuAla 380
 Db 1012 GAAAGCAAGAGAGAGAGAGAGAGAGTACAGGACAGCTACAGAGATGATCCAAATTTGGCC 1071
 Qy 381 ArgTyrlsProValLysProGlnProTyrlsGluGlnMetArgIleHisAlaGluValSer 400
 Db 1072 CGGTATCCGCTAAAGCCACAACTATGAAGAGCAATGCGATCCAGCTCAGCTGAAGTGTCC 1131
 Qy 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
 Db 1132 CGGGCAGCAGCAGAGAGAGAGAGAGAGTGTGATGTTTCTTGGCAATATGCTGAACATAGAAT 1191
 Qy 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
 Db 1192 TCCAGAGATTCCTGCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1251
 Qy 441 ArgAlaAlaMetGluAsnGlnArgSerTyrlsSerMetGluArgThrArgGluAlaGlnGly 460
 Db 1252 AGAGCTGCAATGAGAAACAGCAGCGATGCTATTCATGCAAGAGAGAGAGAGAGAGAGAG 1311
 Qy 461 GlnSerSerTyrlsProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
 Db 1312 CAAAGTCTTATCCAAAG 1371
 Qy 481 IleProLeuAspArgProAspMetArgAlaAspSerLeuArgLysGlnHisHisLeu 500
 Db 1372 ATACCGCTCATGATAGCAGAGCTGAGCGGTGCCATCTCCCTACGAAACACACACCTTA 1431
 Qy 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
 Db 1432 GATCCCGAGCTCTCTGTAGGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1491
 Qy 521 AspSerLeuSerSerSerAspGlnSerGluSerValArgProProProProProProHisLys 540
 Db 1492 GATTCTTGTAGTTGAGACAGAGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1551
 Qy 541 SerLysLysGlyGlyLysMetArgGlnValSerSerSerSerGluGluGluLeuAla 560

Db 1552 TCCAGAAAGGAGGTAATAATGCCAGAGTTTCACTGACAGCTCAGAGGAGGTGGCA 1611
Qy SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
Db 1612 TCCAGCCTGAGTATACAAAGCTGTGATGACGCGGAGATGAAAGCGAGAGGCTAGGTGAG 1671
Qy LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
Db 1672 AAAGAGACATGGAG----- 1686
Qy AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
Db 1687 -----TACAGCTGGTGGAG 1701
Qy AspLeuGluTyrSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr 640
Db 1702 CATCGCTCTGGCAT----- 1716
Qy ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTyrGlnProSerLys 660
Db 1717 -----AGCAGTGGGCGATCCCAATGCTTTTGCACCTGTGACCTGGCGAGCCATCCAAA 1770
Qy AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 1771 GATGAGATCGCTAATTTGGTCGTATTTTATTAATAAGCGTTTAAAGATGGGAGTGA 1830
Qy ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer 700
Db 1831 CCTCGAGATTCAGGAGCAATGCTGGCTTAAAGGTTTAAAGGTTAGGAGGAAAGTCACTCAATCA 1890
Qy GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
Db 1891 GGTGACATTTGGCATTTATTAACCAAGTGAAGGAAAGGAGTTTACGTGATCTAGGA 1950
Qy HisLeuArgProGlyAspGluValLeuGluTyrAsnGlyArgLeuLeuGlnGlyAlaThr 740
Db 1951 CATCTTAGACAGGTGATGAAGTCTTGGATGGATGGAGCTATTTCAGAGGAGCCACA 2010
Qy PheGluGluValTyrAsnIleLeuLeuGluSerLysProGluProGlnValGluLeuVal 760
Db 2011 TTTGAGGAAGTTTACAACATTTCTAGATCCAGCCTGAACCAACCAAGTTGAGTGTGT 2070
Qy ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
Db 2071 GTTTCAGGCCCATTTGGAGATATGCTAGAAATACCTGATAGACATCATGCACTGCAACATGGAA 2130
Qy SerSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
Db 2131 TCCAGTTCTAGTCTCAATTTGAATCTCAAAAATGGATCGTCCGTCTATATCGGTACCTCT 2190
Qy ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
Db 2191 CCCATGAGTCTGGCATGCTGAGGAGATGCCACAGTCTTATCTGGAAGCTTTCAATA 2250
Qy LysLeuTyrPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
Db 2251 AAACATGTTTGGACAGGTTGGTCCACAGTTAATAGTTTACAAATTTTGGAGCAAGGAT 2310
Qy LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
Db 2311 CTCCCTTCCAGGAAGATGGAGGCCAAGAAATCTTATGTTAAATTTACTTCTCTCCA 2370
Qy AspArgSerAspLysAsnLysArgThrLysThrValLysLysThrLeuGluProLys 880
Db 2371 GACAGAAGTGATAAAACAAGAGAGAGAAACAAAACAGTCAAGAAACCTTTGGAAACCCAAA 2430
Qy TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
Db 2431 TGGAAACAGACTTCAATTTATTTCTCTGTTCCAGGAGAAATTCGCGGAACGAATGCTC 2490
Qy GluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu 920

Db 2491 GAAATCACCCCTTTGGGATCAAGCTCGAGTTCGAGAGGAGAAAGTGAATTCCTAGGAGAG 2550
Qy IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisIleTyrTyrLysLeuGln 940
Db 2551 ATTTTAAATGAATTTGGAAACAGCTCTGTAGATGATGAACACACACTGGTCAAACTTCAG 2610
Qy ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgGln 960
Db 2611 ACCCATGATGCTCTCTCATTTGCCACTCTCTCACCTTCTCCATATATGCCAGAGACAG 2670
Qy LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
Db 2671 CTGCATGAGAGAGCCCAACACAGCAGGTTCGAAAGGTCGAAGAGATAAGTGCAGAGCAA 2730
Qy ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
Db 2731 GTATCCCACTATGATCGAGGACGCGGTGGAGTCTGTGTGATATACCGACAGATGGC 2790
Qy ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
Db 2791 CGTGATCTTCAAGCTTCAATATCATGTCGACAGACAGATCATGTCTCATCAACCATTCG 2850
Qy SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTyrSerPro 1040
Db 2851 TCACCATCAGCTCTCTCATCGAGTAGATGTTATAGAAAGACTAGTCTCATGTGCGCT 2910
Qy SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
Db 2911 AGTGTCCCTCTCTCTCAAGGAATGTGGAACAGGGGCTTCGAGGAGACACGTGCTACTGGC 2970
Qy HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
Db 2971 CATTCACATACATTAAGCGGATGGACAGACACCGTGTCTCATGACGACCATTACTCTCA 3030
Qy AspArgAspArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThr 1100
Db 3031 GAGAGAGACAGGATTTGTAAGCAGCAGATAGACAGCCATATCAGATCCAGATCAACA 3090
Qy GluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArgProAsp 1120
Db 3091 GAACAACGGCTCTCTAGAGCGGACCAACCGCTCCAGATCTCTGAAACGTGCTGAT 3150
Qy ThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerPro 1140
Db 3151 ACAACCTCATGAGTGTGATGCTTCAATTAAGTCTGAAGATCTGCCCCCTCTTCACT 3210
Qy AlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrPro 1160
Db 3211 GCCTTATCGAGTCTCACCTCGCATGGCTCTGTCCAGACAGTCCGCTCAAGTACTCCC 3270
Qy GlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArg 1180
Db 3271 GTAAACAGAGCAAGGGCGGCGAGCTTCTCAGCTCCCAAGGGGAACTGGAAGA 3330
Qy SerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnVal 1200
Db 3330 ----- 3330
Qy AlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTyrPhe 1220
Db 3330 ----- 3330
Qy ProHisArgGlyAlaAspThrValSerThrLysSerAspSerAspValSerAspVal 1240
Db 3330 ----- 3330
Qy SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerVal 1260
Db 3330 ----- 3330
Qy GlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsn 1280
Db 3330 ----- 3330

Qy 1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerLysSerGlyAspMet 1300
 Db 3330 ----- 3330
 Qy 1301 CysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThr 1320
 Db 3330 ----- 3330
 Qy 1321 SerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSer 1340
 Db 3330 ----- 3330
 Qy 1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyGlyLysLeuArg 1360
 Db 3331 ----- 3331
 Qy 1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
 Db 3352 AGCACTGTCCAGAGAGACCGGAGACCGGCTAGCAGTGGAGATGAGAACTGGATGACA 3411
 Qy 1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsn 1400
 Db 3412 AGACAGGCCACCGGGGATCCAGCATGGCAGCATGAACTTACAGCTCAGAGAGAAAT 3471
 Qy 1401 LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly 1420
 Db 3472 CTGATCTTCTGGGTCGGCTGGCTCTGACAGCCAGTTCAGCGATTTCTTGATGTT 3531
 Qy 1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440
 Db 3532 CTGGCCCTGTCTAGTGTGGAGCGCCAGACCTCTGGCGACTCTCTGCAATGGGTGACATT 3591
 Qy 1441 GlnValGlyMetMetAspLysGlyGlnLeuGluValGluIleIleArgAlaArgGly 1460
 Db 3592 CAGTGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3651
 Qy 1461 LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu 1480
 Db 3652 CTGTGTAGTAAACAGGTTTCCAGACATCTGCAGCACCATATGTTCAGGTGATCTGTTA 3711
 Qy 1481 AspAsnGlyValCysIleAlaLysLysLysLysLysLysLysLysLysLysLysLysLys 1500
 Db 3712 GACACGAGAGTCTGCATGACCAAAAGAAACCAAGGTGGCGAGAGAGAGAGAGAGAGAG 3771
 Qy 1501 LeuTyrGlnGlnLeuLeuSerPheGluGlnSerProGlnGlyArgValLeuGlnIleIle 1520
 Db 3772 CTGTATCAGCAGCTGTATCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3831
 Qy 1521 ValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu 1540
 Db 3832 GTCTGGGGAGATTATGGCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 3891
 Qy 1541 LeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPheProSer 1560
 Db 3892 TTAGATGAATGGAATCAACATCAACATCAACATCAACATCAACATCAACATCAACATCA 3951
 Qy 1561 SerLeuValAspProThrSerAlaProLeuThrArgAlaSerGlnSerSerLeuGlu 1580
 Db 3952 TCCCTAGTAGATTCACACCTTGGACCTCTGACAGAGAGAGTTCACCAATCGTCTCTGGA 4011
 Qy 1581 SerSerThrGlyProSerTyrSerArgSer 1590
 Db 4012 AGTTCTACTGGACCTCTTACTCTCTGTTCA 4041

RESULT 9

AF199323
 LOCUS AF199323 5031 bp mRNA linear ROD 04-JUL-2000
 DEFINITION Rattus norvegicus RIM2-2A (Rim2-2A) mRNA, complete cds.
 ACCESSION AF199323
 VERSION AF199323.1 GI:8925859
 KEYWORDS Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

1 (bases 1 to 5031)
 Wang, Y., Sugita, S. and Sudhof, T.C.
 The RIM/MIM family of neuronal C2 domain proteins. Interactions
 with Rab3 and a new class of Src homology 3 domain proteins
 J. Biol. Chem. 275 (26), 20033-20044 (2000)

JOURNAL

20347919

MEDLINE

10748113

PUBMED

REFERENCE

2 (bases 1 to 5031)

AUTHORS

Wang, Y. and Sudhof, T.C.

TITLE

Direct Submission

JOURNAL

University (27-OCT-1999)

Center for Basic Neuroscience, The

University of Texas Southwestern Medical Center, 6000 Harry Hines

Blvd., Dallas, TX 75235-9111, USA

FEATURES

Location/Qualifiers

1..5031

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

1..5031

/gene="Rim2-2A"

52..4110

/codon_start=1

/product="RIM2-2A"

/protein_id="AAF81645.1"

/db_xref="GI:8925860"

/translation="MSAPLGRGRPAPTPAASQPPQPEMPDLSHLTTEERKIIQAVM
 DROKREKESQVLKCLHQOFYKQVKGESQOQKQDGAFTCGICHTKTFAD
 CGHCNCCCTKFCARCGSVLSRNVKVCNLCRQOILTKSGAFWFSNNTQ
 QPQKALGLRSEAPQEKALHEQTFQFGPDSSVPAVERAGHLTRQDSIKG
 SMKHQIASMPDRKRSVSRDQNRKYDQSEREYSQTVPSDSTSPSPSYAR
 RSQEPQFPEPDLNTRDSNRHSKEYTVDDVEDSRDEYRQRRESEYQARY
 SDNLARFYKQFTEQRIHAESRAREHRSYVLSANAELEDSRISLMDRPS
 RORSERRAAMENORSYMERTRAQSSYPORTNHSPTVPRRSPILDRPELAR
 ADSLRKQHLDPSSAVKTKRQMETMLRNDLSQDSVSPVPPPPHKKGKGR
 QVLSSESELASTPEVTSDDVIESVGEKDMESYMLERHASSHSSPMSLHP
 VTQPSKQDRLIGRILINKLKDGSVPRSGAMGLKLVGGKNTSEGLCAFTTKVK
 KSLADTVGHLRDEVLWNGRLLOGATFEVNTILSKPEQVLELVSRPTGDMF
 RIFDTHAQLESSESPSQMDRPSISVTPMSPGMLRDPQFQSLQSLTKLWDFVY
 GHOLIVTILAKOLPSREDGRPNPYKIVFLPDRSKKRTNTVKTLEPKWNTF
 YVPSHREPERFEMLEITLMDQARVRESEFELEILETALLDDEPHYKLTQTD
 VSLPLPHSPYPRRQLHGESPTRLQSKRSISDSVSDYDCEDGVGVSDYKDR
 DLQSTLUSVEQVNSNHSPSGPHRDVIGTRSWSPVPPQPVNVEGLRTRAT
 GHNTISRMDRHVMDDHYSSESDSHFTLTPRSHRQTSSEHHRDGRDCAADRPH
 RSRSTEQRLPILRTTTSRSEADTNLMRSPSLMTGRSAPSPSPALSRSHPTGSQ
 TSPSTPTVTRGRQLPOLPKGTLRGGGKCLRSYVORSTETCLAVEMRNMTRQAS
 RESTDGSMNYSYSGENLIFFQVRLASDQSPDLGLGPAQLVCRQTLATPMGDIQV
 GMDKKQLSEYIIRARGLVVKPGSKTLPAPYKVLKNGVCLAKKTKVAKTLP
 LYQQLSFSESPQGVLIQVWGDIYGRMDHKSFMGVAQIILDELSNMVIGWPKLFP
 PSSIVDPTLAPLTRASQSSLESSTGPSRS"

BASE COUNT 1503 a 1166 c 1246 g 1116 t
 ORIGIN

Alignment Scores:

Pred. No.: 7,72e-197 Length: 5031
 Score: 6522.00 Matches: 1283
 Percent Similarity: 80.71% Conservative: 18
 Best Local Similarity: 79.59% Mismatches: 29
 Query Match: 78.72% Indels: 282
 DB: 10 Gaps: 5

US-09-617-099B-1 (1-1590) x AF199323 (1-5031)

Qy 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
 Db 52 ATGTCGGCTCACTCGGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCTCAGCCT 111
 Qy 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysIleIle 40

761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
2071 GTTTCAGGCGGATGAGATATGCTAGATACCTGATAGCACTCATGACACTGGAA 2130
781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
2131 TCCAGTTCTAGCTCATTTGAATCTCAAAAAATGATCGTCCGCTATATCGGTACCTCT 2190
801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle 820
2191 CCATGAGTCCCTGCACTGAGGAGATGCCACAGTTCATTATCTGGACAGCTTCAATA 2250
821 LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp 840
2251 AACTATGTTTGAAGGTTGGTCCACAGTTAATAGTTACATTTTGGAGCAGAGAT 2310
841 LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro 860
2311 CTCCCTTCCAGGAAGATGGAGCCCAAGGAATCTTTATGTTAAATTTACTTCTTCCA 2370
861 AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys 880
2371 GACAGAGTATATAAAACAAGAGAGAGAACAAAAACAGTCAAGAAAACTTTGGAAACCCAAA 2430
881 TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu 900
2431 TGGNACCAAGCTTTCATTTATCTCCGTTCACCAAGAGAGATTCGGGACGATGCTC 2490
901 GluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu 920
2491 GAAATCACTCTTGGGATCAAGCTCGAGTTCGAGAGGAAGAAAGTGAATTCCTAGGAGAG 2550
921 IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGln 940
2551 ATTATTAATGAAATGGAACAGCTCTCTTAGATGATGAACACACACTGGTACAACTTCAG 2610
941 ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln 960
2611 ACCATGATGTCCTCATTTGCCACTCCCTCACCTCTTCATATATGCCAGGAGACAG 2670
961 LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu 980
2671 CTGCATGGAGAGCCCAACAGCAGGTTGCAAGGTCGAAGAGAAATAGTGACAGCGAA 2730
981 ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly 1000
2731 GTATCCGACTATGACTCGAGGAGCGGCGTGGAGTCTGTGATTAACGACAGATGGC 2790
1001 ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys 1020
2791 CGTGATCTTCAAGCTCCATATATCATGTCGACAGCAAGTCATGTCAAAACCATTCG 2850
1021 SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro 1040
2851 TCACCATCAGCTCTCTCATCGATAGATGTTATAGGAGGACTAGTCTATGTCCTCT 2910
1041 SerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly 1060
2911 AGTGTCCCTCTCTCAAGGAATGTGAAACAGGGGCTTCGAGGACACGCTGCTACTGGC 2970
1061 HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer 1080
2971 CATTAACATCAATTAGCCGATGGACAGACACCGTGTCTATGGACGACCATTACTTCA 3030
1081 AspArgAsp----- 1083
3031 GAGAGACAGTCAATTTCTCACTCTACCTCGCTCTCGACACAGGACAGCAGTGCAT 3090
1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg 1098
3091 CACCACAGGAGATGAAGGATTTGTAAGCAGCAGATAGACAGCCATATCATCAGATCCAGA 3150
1099 SerThrGluGlnArgProLeuLeuGluArgThrThrArgSerArgSerGluArg 1118

3151 TCACAGAACACCGGCTCTCTAGAGCGGACCAACCCGCTCCAGATCTCTGACGT 3210
1119 ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro 1138
3211 GCTGATACAAACCTCATGAGGTGATGATCCCTTCAATTAATGACTGGAAGATCTGCCCTCT 3270
1139 SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer 1158
3271 TCACCTGCGCTTATCGAGGTCTCACCTCGACCTGGCTCTGTCCAGACCACTCGTCAAGT 3330
1159 ThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProLysGlyThrLeu 1178
3331 ACTCCGTAACAGACAGAGGGGCCGAGCTTCTCAGCTCCACCAAGGGAACACTG 3390
1179 GluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLys 1198
3391 GAAAGA----- 3396
1199 GlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGly 1218
3396 ----- 3396
1219 TrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSer 1238
3396 ----- 3396
1239 AspValSerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMet 1258
3396 ----- 3396
1259 SerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMet 1278
3396 ----- 3396
1279 GlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGly 1298
3396 ----- 3396
1299 AspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeu 1318
3396 ----- 3396
1319 GlyThrSerGlyLysLysArgArgSerSerIleGlyAlaLysMetValAlaIleValGly 1338
3396 ----- 3396
1339 LeuSerArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLys 1358
3397 -----GAGGAGGTAAAG 3411
1359 LeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrp 1378
3412 CTACGACGACACTCTCCAGAGAGACAGGAGCCGGCTAGCAGTGGAGATGAGAACTGG 3471
1379 MetThrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGlu 1398
3472 ATGACAGACAGCCCGCCGGAATCCACGATGCGCAGCATGACAGTTACAGCTCAGAA 3531
1399 GlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeu 1418
3532 GGAATCTGATCTTCTCGGGTCCGCTCTGACAGCCAGTTCACGAGATTCCTT 3591
1419 AspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly 1438
3592 GATGTCCTGGCCCTGCTCAGCTAGTGGAGCGCAGACCTCGGCTCTCTGCAATGGGT 3651
1439 AspIleGlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleLeuArgAla 1458
3652 GACATTCAGTGGGATGATGATAAGAGGAGCAGCTGAGGTAGAGATCATCCGGCT 3711
1459 ArgGlyLeuValLysProGlySerLysThrLeuProAlaProTyrValLysValTyr 1478

```

Db      3712  CGAGGCTTGTTAGTAAACAGGTTCCAAAGACACTGCCAGCACCATATGTCAGGTTGAC 3771
QY      1479  LeuLeuAspAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeu 1498
Db      3772  CTGTTAGACACGAGTGTGCATAGCAAAAGAAACCAAGTGGCGAGAAAGACGCTG 3831
QY      1499  GluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGln 1518
Db      3832  GAGCCCTGTATCAGCAGCTGTATCTTTCGAGAGAGCCCAAGGAGAGTGTACAG 3891
QY      1519  IleIleValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGln 1538
Db      3892  ATCAATGTCTGGGAGATTATGCGCATGATCAAAATCCTTTATGGAGTGGCCAG 3951
QY      1539  IleLeuLeuAspGluLeuLeuSerAsnMetValIleGlyTyrPheLysLeuPhePro 1558
Db      3952  ATACTCTTAGATGAACCTGAACCTATCAACATGGTAATGGATGTTCAAACTCTCCCT 4011
QY      1559  ProSerSerLeuValAspProThrSerAlaProLeuThrArgAlaSerGlnSerSer 1578
Db      4012  CTTTCTCTCTAGTAGATCAACCTTGGCACTCTGACAGAGAGCTTCCCAATCGTCT 4071
QY      1579  LeuGluSerSerThrGlyProSerTyrSerArgSer 1590
Db      4072  CTGGAAGTCTACTGGACCTCTTACTCTCGTTCA 4107

```

RESULT 10

```

AF199326
LOCUS      AF199326          5031 bp  mRNA  linear  ROD 04-JUL-2000
DEFINITION Rattus norvegicus RIM2-3B (Rim2-3B) mRNA, complete cds.
ACCESSION  AF199326
VERSION    AF199326.1 GI:8925865
KEYWORDS
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 5031)
            Wang, Y., Sugita, S. and Sudhof, T.C.
            The RIM/NIM family of neuronal C2 domain proteins. Interactions
            with Rab3 and a new class of Src homology 3 domain proteins
            J. Biol. Chem. 275 (26), 20033-20044 (2000)
JOURNAL    20347919
MEDLINE    10748113
PUBMED
REFERENCE  2 (bases 1 to 5031)
            Wang, Y. and Sudhof, T.C.
            Direct Submission
            Submitted (27-OCT-1999) Center for Basic Neuroscience, The
            University of Texas Southwestern Medical Center, 6000 Harry Hines
            Blvd., Dallas, TX 75235-9111, USA

```

FEATURES

```

source
1..5031
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
1..5031
/gene="Rim2-3B"
52..4110
/gene="Rim2-3B"
/codon_start=1
/product="RIM2-3B"
/protein_id="AAF81648.1"
/db_xref="GI:8925866"
/translation="MSAPLGPRGRPAPTPAASQPPPPQPMPLSHLTBEERKIIQAVM
DRQKEEKEQSVLKKLHQFEMVKEQVKEESQOQEQKGDAPTGGICHTKTFAD
GGHNCVQCTKEFACGGSVLSKNVWVNCNCRKQQLITKSGAPYNGSTTPQ
QPDQALRGSRSEAPQERAKLHEQTQGGPGSSVPAVERGARHGLTRQDSING
SGKKHQADMSFDRKRSVSDQNRRYDQSEEREYSQYPSFSTWPSPSYADR
RSQRFQFYFVPEHLNRRNRHRSKEYIVDDVEDSRDEYERQREERYQARYR
SDPNLARYPVKQPYEQMRIMAEVSRARHRRHSVLSNAELSDRSISLMDRPS
RORSVSEREAAMENORSYSMERTREAGQSSYPQRTNHSPTPRSPITLDRPELR
AUSLRKQHLDPSAVRKTREKMETMLRNDLSLSDQSESVRPPPPRPHKSKKGKMR

```

gene

CDS

```

QVLSSEELASTPEVTSDDVEIESGVEKGMVEYSVLEHSMHSSEASPMSLHP
VTWPSKXDELIGRIILNKLKDGSPROSGAMLGKVGKMTESGRICAFITKVK
KGLADTVGLHLPDGEVLEWNGRLLOCATFEVYNIILSPFQVQLVLSRPTGDMF
RIPDTTQALLESSSFESQMDRPSISVTPSPGMLRDVPQFLSGQLSKMFPDKV
GHQIVITLGAKDLPSREDGPRPYKIFYLPSRDKNRTKTKTKLPEKWNQTF
IYSPHLRFRERMLEITLWDQAVRBEESFLAGEILIETALTALDDPHWYKLTQHD
VSSPLPHSPMPRROLHGSPTRRLQSRKRIISDSVSDYDCSDGVGVSDYRHDGR
DLQSTLSVPPQVMSNHCSPSGSPHVDVIGTRTSPSPSPVPQRNVQGLCTRAT
GHYNTISRMDRHRYMDHYSERASHFTLPRSRHRTSEHHHRDGRDCEAADRQPHY
RSRTEQRLLERTITRSERADTNLMRSPSLMTGSRAPPSPALSRPTGTSVQ
TSPSTPTVTRGRQLPQPKGLERGGKLRSTVQRSTETGLAVEMRNWMTQAS
RESTDGSNYSSEGNIIFQVRLASQFSDFLDGLGPAQLVGRQTLATPAGMDIQV
GMWDKGLQLEVIIRAGLVVPGSKTLPAPYKVLVLDNGVCTAKKTKVAKTLEP
LYQQLSFEESPQKVLQIVVPGDYGRMDHKSFMGVAQLILDELLELSNMVGNPKLFP
PSSLVDPLAPLTPRASQSSLSSTGTSYRS"
BASE COUNT 1503 a 1166 c 1246 g 1116 t
ORIGIN
Alignment Scores:
Pred. No.: 7,72e-197 Length: 5031
Score: 6522.00 Matches: 1283
Percent Similarity: 80.71% Conservative: 18
Best Local Similarity: 79.59% Mismatches: 29
Query Match: 78.72% Indels: 282
DB: 10 Gaps: 5
US-09-617-099B-1 (1-1590) x AF199326 (1-5031)
QY 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
Db 52 ATGTGGCTCCACTCGGGCCCGGGCGCGCGCTCCACCCGGCGGCTCTCAGCCT 111
QY 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysIleLe 40
Db 112 CTCTCGAGCCGAGATCCGAGCTCAGCCACTCAGCGAAGAGAGGAGGAAATCATC 171
QY 41 LeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeuLysIle 60
Db 172 CAGGCTGTCATGATGTCGTCAGAGAAAGAGAGAGGAGGAGGAGCTCGTCTCAAA --- 228
QY 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
Db 228 ----- 228
QY 81 ValAsnAsnValLeuGlnProGlnLysGlnProAsnGluLysGluProGlnThrLys 100
Db 229 -----AAG 231
QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120
Db 232 CTGCATCAACATTTGAATGTTATAGGAGCAAGTCAAGAAATGCGGAGAGGAATCACA 291
QY 121 GlnGlnGlnGluGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
Db 292 CAGCAGCAAGAGCAGAGGCGGAGCGCCCGACCTCGCGCATCTGCCACACAGCAAAATTT 351
QY 141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
Db 352 GCAGATGGATGGCGCCATAACTGTTTATATGCCAACCAAGATCTCTGCTGCTGTGGA 411
QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
Db 412 GGTGAGGTGTATATACGCTCAACAGGTATATGCGGTGTGTATATTTGTCGCGGAAAAA 471
QY 181 GlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
Db 472 CAAGAAATCTCACTAAGTCCGGCGCGTGTGTTTACATATAGTGGTCTAATACACCGAG 531
QY 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluAlaProGlnGlnLysLys 220
Db 532 CAGCCTGATCAAAAGGCTCTTCGAGGGCTTCGAGTGGAGGAGGAGGAGGAGGAGGAG 591
QY 221 AlaLysLeuHisGluGlnProGlnPheGlnAlaProGlyAspLeuSerValProAla 240

```

Db	592	GCACAACTGATGAGCAGACGACGATCCAGGACCCCGGCTGCTCATCAGTACCTGCA	651	1672	AAAGGAGACATGAG	-----	1686
Qy	241	ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly	260	601	AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyHisSerLeuGluGlu	620	1687
Db	652	GTGAGAGAGGTGAGCTCATGGGCTCACAAGACAGGATTCATTATTAAGATGCTCAGGA	711	1687	-----	TACAGCTGGTGGAG	1701
Qy	261	ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer	280	621	AspLeuGluTyrSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr	640	1702
Db	712	ATGAAGCACCAGATCCGAGTGCATGCTTCACAGCAAAAGAGATCGCTCAGTGTCC	771	1702	CATGGCTCTGGCAT	-----	1716
Qy	281	ArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgLysArgTyrSerGlnTyrVal	300	641	ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTyrGlnProSerLys	660	1717
Db	772	AGGGATCAAAATCGAAGATACGACCAAGAGTGAAGAGAGAGAGATATTACAGTATGT	831	1717	-----	AGCAGTGAGGCATCCCAATGCTCTTCCACCTGTGACCTGGCAGCATCCAAA	1770
Qy	301	ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg	320	661	AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLysAspGlySerVal	680	1771
Db	832	CTTTCAGATAGCAATGCTCAGATCTCCATCAGATTATGCTGATAGCATCTCAACGT	891	1771	GATGGAGATCGGCTTAATGGTGTATTTTAAATAAGCTTAAAGATGGAGTGTA	1830	1771
Qy	321	GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly	340	681	ProArgAspSerGlyValaMetLeuGlyLeuLysValValGlyGlyMetThrGluSer	700	1831
Db	892	GAACCTCAATTTATGAGAGACCTGATCATTTAAATATAGGATTTCAACGAGAGGC	951	1831	CCTCGAGATTCAGGACCAATGCTGGCTTAAAGGTTGTAGGAGGAAAGATGACTGAATCA	1890	1831
Qy	341	HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr	360	701	GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly	720	1891
Db	952	CATAGACATTCACAAAGATATATTGTAGACGACGAAGATGTGAGAGCAGAGATGAAT	1011	1891	GGTCGACTTTGGCATTTATTACCAAGTGAAAGGAGAGTTTACTGATGACTGAGGA	1950	1891
Qy	361	GluArgGlnArgGluGluGluTyrGlnAlaArgTyrArgSerAspProLeuLeuAla	380	721	HisLeuArgProGlyAspGluValLeuGluLysAsnGlyArgLeuGlnGlyAlaThr	740	1951
Db	1012	GAAAGCAAGAGAGAGAGGAGTACCGACGCTCAGAAAGTGCATCAAAATTTGGCC	1071	1951	CATCTTAGACACGAGTATGGAAGTCTTGAATGCAATGGAAGCTATTGCAAGGACCA	2010	1951
Qy	381	ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer	400	741	PheGluGluValTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuVal	760	2011
Db	1072	CGGTATCCGGTAAGCCACACCCATATGAGACCAATGCGGATCCACGCTAAGTGTCC	1131	2011	TTTGAGGAGTTTACCAATTTCTAGAATCCAAAGCTTGAACCAAGCTTGAAGCTGTT	2070	2011
Qy	401	ArgAlaArgHisGluArgHisSerAspValSerLeuAlaAsnAlaGluLeuAsp	420	761	ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu	780	2071
Db	1132	CGGGCACGACACGAGAGAGGACAGTGTATTTCTTGGCAATGCTGAATAGAGAT	1191	2071	GTTCACAGCCGATTTGGAGATATGCTAGATATCTGATAGCATTCTGCACACTGGAA	2130	2071
Qy	421	SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg	440	781	SerSerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer	800	2131
Db	1192	TCAGGATTTCTGTGTAAGGATGATAGACCATCAAGGCAAGATCTGTATCTGAACGT	1251	2131	TCAGTCTTAGCTCATTTGAATCTCAAAATGGAATGCTGCTCTATATCGGTACTCT	2190	2131
Qy	441	ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgLysGlnGly	460	801	ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle	820	2191
Db	1252	AGAGCTGCATGAAACAGGCGATCGTATTCAATGGAAGAACTCGAAGCTCAGGA	1311	2191	CCCATGAGTCTCGGATGCTGAGGAGTCTCCACAGTCTTATCTGCAAGCTTCAATA	2250	2191
Qy	461	GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro	480	821	LysLeuTyrPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyValLysAsp	840	2251
Db	1312	CAAGTTCCTATCCACAAAGGACCAAAATCATAGTCTCTTACCCCTCGCAGGACCT	1371	2251	AAACTATGTTTGACAGGTTGGTCCAGTTATAGTTACAAATTTGGGCAAGGAT	2310	2251
Qy	481	IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLysLeu	500	841	LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro	860	2311
Db	1372	ATACCGCTCGATAGACGAGTGGCGGCTGCGGATCTCCCTACGGAACACACCACTTA	1431	2311	CTCCCTTCAGGAGAGATGGGAGGCAAGGAATCTTATGTTAAATTTACTTCTTCCA	2370	2311
Qy	501	AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn	520	861	AspArgSerAspLysAsnLysArgThrLysThrValLysLysLysLysLysLysLys	880	2371
Db	1432	GATCCAGCTCTGCTGTAGGAAACGAGCGGAGGAAAAATGGAACAAATGTGAGGAT	1491	2371	GACAGAGTGATTAATAACACAGAGAGAAACAAACAGTCAAGAAATCTTGGAAACCCAA	2430	2371
Qy	521	AspSerLeuSerSerAspGlnSerGluSerValArgProProProProProHisLys	540	881	TyrAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluMetLeu	900	2431
Db	1492	GATCTTTGAGTTTCAGACGATCTGAGTTCAGTGGGCGCCCGCCCGGCTCATAAA	1551	2431	TGGAAACAGACTTTCATTTATTTCTCTCTTCCCGAGAGAAATCCGGGAGCAATGCTC	2490	2431
Qy	541	SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAla	560	901	GluIleThrLeuTyrAspGlnAlaArgValArgGluGluGluSerGlnPheLeuGlyGlu	920	2491
Db	1552	TCCAGAAAGAGGAGTAAATGCGCCAGGTTTCACTGAGCAGCTCAGAGGAGGATGGCA	1611	2491	GAATCAACCTTTGGGATCAAGCTCGAGTTCGAGAGGAGAAAGTGAATTTCTTAGAGAG	2550	2491
Qy	561	SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu	580	921	IleLeuIleGluLeuGluThrAlaIleLeuAspAspGluProHisTyrTyrLysLeuGln	940	2551
Db	1612	TCCAGCCCTGATATACAGCTGTGATGACGTGGAGATTTGAAAGCGAGAGCGTATGAG	1671	941	ThrHisAspValSerLeuProLeuProArgProSerProTyrLeuProArgArgGln	960	2551
Qy	581	LysGlyAspSerGlnLysGlyLysArgLysThrSerSerGluGlnGlyValLeuSerAspSer	600	2611	ACCATGATGCTCTCTTCCCTCCCTCACCCTCTCCATATATGCCACGGAGACAG	2670	2611

[illegible]

Qy	421	Ser	Arg	Ile	Ser	Leu	Leu	Arg	Met	Asp	Arg	Pro	Ser	Arg	Gln	Arg	Ser	Val	Ser	Glu	Arg	440
Db	1192	TCC	AGG	AT	T	T	C	T	C	T	G	C	T	A	A	G	A	T	A	G	A	1251
Qy	441	Arg	Ala	Ala	Met	Glu	Asn	Gln	Arg	Ser	Tyr	Ser	Met	Glu	Arg	Thr	Arg	Glu	Ala	Gln	Gly	460
Db	1252	AG	ACT	G	C	A	T	G	A	A	A	C	C	A	G	A	T	C	G	T	A	1311
Qy	461	Gln	Ser	Ser	Tyr	Pro	Gln	Arg	Thr	Ser	Asn	His	Ser	Pro	Pro	Thr	Pro	Arg	Arg	Ser	Pro	480
Db	1312	CAA	AGT	T	T	C	T	T	A	C	A	A	A	G	A	C	A	A	T	C	T	1371
Qy	481	Ile	Pro	Leu	Asp	Arg	Pro	Asp	Met	Arg	Arg	Ala	Asp	Ser	Leu	Arg	Lys	Gln	His	His	Leu	500
Db	1372	AT	AC	C	G	T	C	G	A	T	A	C	C	A	G	A	G	T	G	C	C	1431
Qy	501	Asp	Pro	Ser	Ser	Ala	Val	Arg	Lys	Thr	Lys	Arg	Glu	Lys	Met	Glu	Thr	Met	Leu	Arg	Asn	520
Db	1432	GAT	CCC	AG	C	T	C	T	G	T	A	G	A	A	A	C	G	A	A	A	A	1491
Qy	521	Asp	Ser	Leu	Ser	Ser	Asp	Gln	Ser	Glu	Ser	Val	Arg	Pro	Pro	Pro	Arg	Pro	His	Lys	540	
Db	1492	GAT	T	C	T	T	T	AG	T	T	C	A	G	A	C	C	A	G	A	C	A	1551
Qy	541	Ser	Lys	Lys	Gly	Gly	Lys	Met	Arg	Gln	Val	Ser	Leu	Ser	Ser	Ser	Ser	Glu	Glu	Leu	Ala	560
Db	1552	TCC	AAA	AAA	AG	AG	GT	AAA	AT	T	G	C	G	C	A	G	T	T	C	A	C	1611
Qy	561	Ser	Thr	Pro	Glu	Tyr	Thr	Ser	Cys	Asp	Asp	Val	Glu	Leu	Glu	Ser	Glu	Ser	Val	Ser	Glu	580
Db	1612	TCC	AC	G	C	T	C	AG	T	A	T	A	C	A	G	T	G	T	A	G	T	1671
Qy	581	Lys	Gly	Asp	Ser	Gln	Lys	Gly	Lys	Arg	Lys	Thr	Ser	Glu	Gln	Gly	Val	Leu	Ser	Asp	Ser	600
Db	1672	AA	G	G	A	C	A	T	G	G	A	G	A	G	A	G	A	G	A	G	A	1686
Qy	601	Asn	Thr	Arg	Ser	Glu	Arg	Gln	Lys	Leu	Arg	Met	Tyr	Tyr	Gly	Gly	His	Ser	Leu	Glu	Glu	620
Db	1687	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1701	
Qy	621	Asp	Leu	Glu	Tyr	Ser	Ser	Glu	Pro	Gln	Ile	Lys	Asp	Ser	Gly	Val	Asp	Thr	Cys	Ser	Thr	640
Db	1702	CAT	CG	T	C	T	T	G	CA	T	---	---	---	---	---	---	---	---	---	---	1716	
Qy	641	Thr	Leu	Asn	Glu	Glu	His	Ser	His	Ser	Asp	Lys	His	Pro	Val	Thr	Trp	Gln	Pro	Ser	Lys	660
Db	1717	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1770	
Qy	661	Asp	Gly	Asp	Arg	Leu	Ile	Gly	Arg	Ile	Leu	Asn	Lys	Arg	Leu	Lys	Asp	Gly	Ser	Val	680	
Db	1771	GAT	G	G	A	T	CG	CT	AA	T	T	GT	GC	T	AT	T	T	TA	TA	AT	A	1830
Qy	681	Pro	Arg	Asp	Ser	Gly	Ala	Met	Leu	Gly	Leu	Lys	Val	Val	Gly	Gly	Lys	Met	Thr	Glu	Ser	700
Db	1831	CCT	CG	A	T	T	C	AG	A	T	T	C	G	G	C	T	T	A	A	G	T	1890
Qy	701	Gly	Arg	Leu	Cys	Ala	Phe	Ile	Thr	Lys	Val	Lys	Lys	Gly	Ser	Leu	Al					

Db	2131	TCAGGTTCTAGCTCATTTCGAATCTCAAAAATGGAATCGTCCGTCTATATCGGTACCTCT	2190
QY	801	ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle	820
Db	2191	CCCATGAGTCTGGCATGCTGAGGAGATGCCACAGTCTTATCTGGACAGCTTTCATA	2250
QY	821	LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp	840
Db	2251	AAACTATGGTTTGACAAAGGTTGGTCACCAAGTTAATAGTTACAAATTTTGGAGCAAAAGGAT	2310
QY	841	LeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuPro	860
Db	2311	CTCCCTTCACGGGAAGATGGGAGCCCAAGGAATCCCTATGTTAAAAATTTACTTCTTCCA	2370
QY	861	AspArgSerAspLysAsnLysArgCysThrLysThrValLysLysThrLeuGluProLys	880
Db	2371	GACAGAGTGTATAAACAAGAGAGAGACAAACAAGTCAAGAAACTTTTGGAAACCCAAA	2430
QY	881	TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu	900
Db	2431	TGAAACAGACTTTCATTTTATTTCTCTCTGTTTCCACGAAGAGAAATTCGGGAACGAATGCTC	2490
QY	901	GluIleThrLeuTrpAspGlnAlaArgValArgGluGluGluSerGluPheLeuGlyGlu	920
Db	2491	GAATCACCCTTTGGGATCAAGCTCCAGTTCGAGAGGAAGAAAGTAGTAATTTCTTAGGAGAG	2550
QY	921	IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGln	940
Db	2551	ATTTTAATTTGAATTTGAAACAGCTCTGTAGATGATGACCAACACCTGGTACAAACTTCAG	2610
QY	941	ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln	960
Db	2611	ACCATGATGTCCTCTCTCATTTGCCACTCCCTCCACCTTCTCCATATATGCCACGAGACAG	2670
QY	961	LeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGlu	980
Db	2671	CTGCATGGAGAGAGCCCAACACACAGGTTTGCAAAGGTCGAAGAGTAAGTGCACGCCAA	2730
QY	981	ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly	1000
Db	2731	GTATCCGATATGACTGCGAGGACGGCGTGGGAGTCGTGTTCAGATTACCGACACGATGGC	2790
QY	1001	ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys	1020
Db	2791	CGTGAATCTTCAAAGCTCCACATTTATCAGTGCCAGAACAGTCAATGTCATCAACACCATTGC	2850
QY	1021	SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro	1040
Db	2851	TCACCATCAGGCTCTCTCATCAGTAGATGTTATAGGAAGGACTAGTTCATGTCGCCT	2910
QY	1041	SerAlaproProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly	1060
Db	2911	AGTGTCCCTCTCTCAAAGGAATGTGAACAGGGGCTTCGAGGGACACGTGTACTGGC	2970
QY	1061	HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer	1080
Db	2971	CATTACAAATACATTAGCCGATGGACAGACACCGTGTCAATGACGACCATTTACTTCTCA	3030
QY	1081	AspArgAsp-----	1083
Db	3031	GAGAGAGACAGTCATTTTCTCACTCTACTCTGCTCTCGACACAGCGACGACCATGAGCAT	3090
QY	1084	-----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg	1098
Db	3091	CACCACAGGAGTGGAGGGGATTTGTAGCGACGACATAGACAGCCATATCACAGATCCAGA	3150
QY	1099	SerThrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArg	1118
Db	3151	TCAACAGAACAAAGGCTCTCTCTAGAGCGGACCAACACCGCTTCAGATCTCTGAAACGT	3210
QY	1119	ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro	1138

Db 3211 GCTGATACAAACCTCATGAGGTCGATGCTTCATTAATGACTGGGAAGATCTGCCCTCCT 3270
 Qy 1139 SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer 1158
 Db 3271 TCACCTGCTTATCGAGGCTCTCACCTCGCACTGGCTCTGTCCAGACAGTCCGTCAGT 3330
 Qy 1159 ThrProGlyThrGlyArgGlyArgGlnLeuProGlnLeuProProGlyThrLeu 1178
 Db 3331 ACTCCGCTAACAGGACGAGGCGCGGAGCTCTCAGCTCCCAACCAAGGGAACACTG 3390
 Qy 1179 GluArgSerAlaMetAspIleGluArgAsnArgGlnMetLysLeuAsnLysTyrLys 1198
 Db 3391 GAAAGA----- 3396
 Qy 1199 GlnValAlaGlySerAspProArgLeuGluAspTyrHisSerLysTyrArgSerdgly 1218
 Db 3396----- 3396
 Qy 1219 TrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspValSer 1238
 Db 3396----- 3396
 Qy 1239 AspValSerAlaValSerArgThrSerSerAlaSerArgPheSerThrSerTyrMet 1258
 Db 3396----- 3396
 Qy 1259 SerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMet 1278
 Db 3396----- 3396
 Qy 1279 GlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerdgly 1298
 Db 3396----- 3396
 Qy 1299 AspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeu 1318
 Db 3396----- 3396
 Qy 1319 GlyThrSerGlyLysArgSerSerIleGlyAlaLysMetValAlaIleValgly 1338
 Db 3396----- 3396
 Qy 1339 LeuSerArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLys 1358
 Db 3397-----GGAGGAGTAAAG 3411
 Qy 1359 LeuArgSerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrp 1378
 Db 3412 CTACGCACTGTCAGAGAAGACGAGACCGGCTAGCAGTGGAGATGAGGAACCTGG 3471
 Qy 1379 MetThrArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGlu 1398
 Db 3472 ATGACACAGCAGCGCCGCGGAATCCAGGATGGCAGCATGAACAGTTACAGCTCAGAA 3531
 Qy 1399 GlyAsnLeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeu 1418
 Db 3532 GGAATCTGATCTTCTCGGCTCGCTGGCTCTGACAGCCAGTTCAGCGATTCTCT 3591
 Qy 1419 AspGlyLeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGly 1438
 Db 3592 GATGCTCTGGGCGCTGCTCAGCTAGTGGAGCGCAGACCTGGCACTCTGCAATGGGT 3651
 Qy 1439 AsplleGlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleAla 1458
 Db 3652 GACATTGAGTGGGATGATGATGAAGAGGAGCAGCTGGAGGTAGATCATCCGGCT 3711
 Qy 1459 ArgGlyLeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyr 1478
 Db 3712 CGAGGCTTGTAGTAAACAGGTTCCAGACACTGCCAGCACCATATGTCAAGTGTAC 3771
 Qy 1479 LeuLeuAspAsnGlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeu 1498
 Db 3772 CTGTTAGACAACGGAGTCTGCATCCCAAAAGAAAAACCAAGGTGGCGAGAAAGCGCTG 3831

Qy 1499 GluProLeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGln 1518
 Db 3832 GAGCCCTGTATCAGCAGCTGTATCTTCGAGAGAGCCCCCAGGGGAGGTGTGTACAG 3891
 Qy 1519 IleIleValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGln 1538
 Db 3892 ATCATTTGTCTGGGAGATTATGGCGCATGATCACAAATCCTTTATGGAGTGGCCAG 3951
 Qy 1539 IleLeuLeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPhePro 1558
 Db 3952 ATATCTTTAGATGAACATGGAACATCAACATGTAATGGTTCGTTCAAACTCTTCCT 4011
 Qy 1559 ProSerSerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSer 1578
 Db 4012 CTTCTCTCCCTAGTAGATCCAACTTGGCACTTTCACAGAGAGCTTCCCAATCGTCT 4071
 Qy 1579 LeuGluSerSerThrGlyProSerTyrSerArgSer 1590
 Db 4072 CTGGAAGATTCTACTGGACCTTCTACTCTCGTTCA 4107
 RESULT 12
 AF199332 LOCUS AF199332 5031 bp mRNA linear ROD 04-JUL-2000
 DEFINITION Rattus norvegicus RIM2-5C (Rim2-5C) mRNA, complete cds.
 ACCESSION AF199332
 VERSION AF199332.1 GI:8925877
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 5031)
 AUTHORS Wang, T., Sugita, S. and Sudhof, T.C.
 TITLE The RIM/NIM family of neuronal C2 domain proteins. Interactions with Rab3 and a new class of Src homology 3 domain proteins
 JOURNAL J. Biol. Chem. 275 (26), 20033-20044 (2000)
 MEDLINE 20347919
 PUBMED 10748113
 REFERENCE 2 (bases 1 to 5031)
 AUTHORS Wang, Y. and Sudhof, T.C.
 TITLE Direct Submission
 JOURNAL Submitted (27-OCT-1999) Center for Basic Neuroscience, The University of Texas Southwestern Medical Center, 6000 Harry Hines Blvd., Dallas, TX 75235-9111, USA
 FEATURES
 source
 1. 5031
 location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 1. 5031
 /gene="Rim2-5C"
 52. 4110
 /codon_start=1
 /product="RIM2-5C"
 /protein_id="AAF81654.1"
 /db_xref="GI:8925878"
 /translation="MSAPLGRPRPAPTAASQPPPEMDLSHTEERKIIQAVM
 DQCKKEQKQSVLKHLHQPEMYKQVKMGESQOQKQDAPTCGICHKTKFAD
 GCGHNCVCTKFCARCVSLNSKVMVNCILCKQRIILTKSGAFVYNSGNTPO
 OPQKALGLSEAPQEKAKLHEQTFQFPDSDSVPAVERGARGLITRODSIKNG
 SMKHQIASDMPDRKSPSVSRDNRKIDQSEHSISQIVPDSSTWPSFSDIAR
 RSQRPEQFPEPDHLYRNRGRHSKTYIVDEDEVRDESDTQRRESEYQAYR
 SDNLARYPVKPYEESQMAHAESVAREHRSVLSANALEDSRIILLRDRS
 RORSVRRRAAMNORSYMERTEAQCSYPTORTNHSPTPRRSPILDRPELR
 ADSLRKOHLDPPSAVKTREKMETMLNDLSLSDQSESVRPPPHKSKGKWR
 QVLSSEELASTPFTSCDDVIESSEVSGKGMETSLKLVGKGMTEGRLCAFTTKVK
 VTWPSKDKORLLIRLLNKLKDGSPVRSGLMLGLKLVGKGMTEGRLCAFTTKVK
 KGSLSADTVHLRPGDEVLWNGRLLOQATFEVNIILEKQPEQVSLVSVPTGDMF
 RLPDSTHAQLSESSSSPESQMDRPSISVTPSPMSGMLRDVPQFLSQLSIKLWFDKY
 GHQLIVTILGADLPSPEDGRPRNPYKVIYFLPDRSDKNKERTNTVKTLEPKNQIF

IYSPVRRERFRMLEITLWDOARVRESESEFLGILILETALDLDDEPHWKLQTHD
 VSSLPLPSPWPRQLHGESPTRELORSKEISDSSEVSDYDCBGVGVSDYRDRG
 DLQSTLSVPEQVMSNHCSPSGSHRVDVIGTRSNPSFVPPORNYEQGLRGTAT
 GHYNTISMDRHRVMDHYSSERDSHEFLTPRSRHKQTSSEHHRDGRDCEAADROPTV
 KRSRTEQRLPRLTRTSRERADTNLMRSMPLTGRSAPPSPALSRSHPTGSGVQ
 TSPSTPTVGRGRQLPPLPPGRLTGERGGKLRSTVORSTETGLAVRNWMTQAS
 RESTDGSNYSSEKMLIPPVRLASDQFSDFLDGLPAQLVGRQTLATPMAGDIQV
 GNMDDKQGLEVILIRAGLVKPGSKTLPAPYKVYLLDNGVICIAKKTIVARKTLEP
 LFCQLSFEESPGKVLIQIWMGDYRMDKSPFVQAQLLLELSELSNWNVGNFKLFP
 PSSLDPTLPTLRASQSSLESSTGSPSRSS*

BASE COUNT 1503 a 1166 c 1246 g 1116 t

ORIGIN

Alignment Scores:

Pred. No.: 7,72e-197 Length: 5031
 Score: 6522.00 Matches: 1283
 Percent Similarity: 80.71% Conservative: 18
 Best Local Similarity: 79.59% Mismatches: 29
 Query Match: 78.72% Indels: 282
 DB: 10 Gaps: 5

US-09-617-099B-1 (1-1590) x AF199332 (1-5031)

QY 1 MetSerAlaProLeuGlyProArgGlyArgProAlaProThrProAlaAlaSerGlnPro 20
 DB 52 ATGTCGGCTCCACTCGGGCCCGGGCCCGCGGCTCCACCCCGCGGCTCTCAGCCT 111
 QY 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluGluArgLysIlelle 40
 DB 112 CCTCGCAGCCGAGATCGGACCTCAGCCACCTCAGGAAAGAGAGAGAGAAATCATC 171
 QY 41 LeuAlaValMetAspArgGlnLysLysGluGluLysGluGlnSerValLeuLysIle 60
 DB 172 CAGGCTGTCTATGATCGTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228
 QY 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyIleThrGluLeu 80
 DB 228 ----- 228
 QY 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
 DB 229 -----AAG 231
 QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120
 DB 232 CTGCATCAACAATTTGAAATGTATAAGGAGCAAGTCAAGAAATGGGAGAGGAATCACAA 291
 QY 121 GlnGlnGlnGlnLysGlyAspAlaProThrCysGlyIleCysHisLysThrLysPhe 140
 DB 292 CAGCAGCAAGAGCAGAGGGCGGCGCCGACCTCGCGCATCTGCCACAGACAAATTT 351
 QY 141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
 DB 352 GCAGATGGATGCGGCCCAACTTTCATATTGCCAAACCAAGTTCTGTGCTCGTTGTGGA 411
 QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
 DB 412 GGTGAGTGTCAATACGCTCAACAAGAGTTATGTGGGTGTGTAATTTGTGCCAAACAA 471
 QY 181 GlnGluIleLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
 DB 472 CAGAAATCCTCAATGTCGGCGGCGGTGTTTACATAGTGGGTCTAATACACCGCAG 531
 QY 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluAlaProGlnGluLysLys 220
 DB 532 CAGCCTGATCAAAAGGCTCTTCGAGGCTTCGCAGTGAGGAGCCCTCAGGAGAGAGAG 591
 QY 221 AlaLysLeuHisGluGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAla 240
 DB 592 GCAAACTGATGAGCAGACGAGTTCACGGGACCCCGGTCAGTCACTCAGTACCTGCA 651
 QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
 DB -----TACAGCTGGTGGAG 1701

DB 652 GTTGAGAGAGCTCGAGCTCATGGCTCACAGACAGGATTTCTATTAGAATGGCTCAGGA 711
 QY 261 VallyHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
 DB 712 ATGAAGCACCAGATCGCCAGTGACATGCTTTACAGACAGAAAAGAGTCCCTCAGTGTC 771
 QY 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrVal 300
 DB 772 AGGATCAAAATCGAAGATAGACCAAGTGAAGAGAGAGAGATATTTCAGATATGTT 831
 QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgArgSerGlnArg 320
 DB 832 CTTTCAGATAGCACAATGCTTAGATCTCCATCAGATTATGCTGATAGACGATCTCAACGT 891
 QY 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340
 DB 892 GAACCTCAATTTTATGAAGAACCTGATCATTTAAATTTATAGGATTTCTAAACAGGAGAGC 951
 QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
 DB 952 CATGACATTTCCAAAGAGTATATTGTAGACGACGAGATGTGGAGCAGAGATGAATAT 1011
 QY 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
 DB 1012 GAAAGACAAAGAGAGAGAGAGTACCCAGGACGCTACAGAGTGTATCCAAATTTGGCC 1071
 QY 381 ArgTyrProValLysProGlnProTyrGluGlnGlnMetArgIleHisAlaGluValSer 400
 DB 1072 CGTATCCGGTAAAGCCCAACCCCTATGAGAGCAAAATGGCGATCCACGCTGAAGTGC 1131
 QY 401 ArgAlaArgHisGluArgArgHisSerAspValSerLeuAlaAsnAlaGluLeuAsp 420
 DB 1132 CGGGCACACAGAGAGAGAGGACAGTGTATTTCTTTGGCAATGCTGAATAGAGAT 1191
 QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
 DB 1192 TCCAGATTTCTCTCTAAGATGATAGACCATCAGGCAAGATCTGTATCTCAACGT 1251
 QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
 DB 1252 AGAGCTGCAATGGAACCAACGAGATCGTATTTCATGGAAGAACTCGAGAGCTCAGGA 1311
 QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProThrProArgArgSerPro 480
 DB 1312 CAAAGTTCTTATCCAAAGGACCAAAATCATAGTCTCTCTTACCCCTCGCAGAGAGCCCT 1371
 QY 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
 DB 1372 ATACCGCTCGATAGACAGAGCTGAGGCGTGCAGCTCCCTACGGAACCAACACCACTTA 1431
 QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
 DB 1432 GATCCAGCTCTGCTGTAGGAAACGAGGAGGAGAAATAATGGAACCAATTTGTAGGAAT 1491
 QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
 DB 1492 GATTTCTTGAGTTTACAGCCAGTCTGAGTCAGTGAGGCGCCGCCCCAGGCTCATAAA 1551
 QY 541 SerLysLysGlyGlyLysMetArgGlnValSerLeuSerSerSerGluGluLeuAla 560
 DB 1552 TCCAGAAAGAGAGGTAATGCGCAGGTTTCACTGAGCAGCTCAGAGAGGAGTTGGCA 1611
 QY 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
 DB 1612 TCCAGCCTCGATATACAGCTGTATGACGTGGAGATTGAAGCGAGAGCGTACGTGAG 1671
 QY 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
 DB 1672 AAAGAGACATGGAG----- 1686
 QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
 DB -----TACAGCTGGTGGAG 1701

Qy	621	AspLeuGluTrpSerGluProGlnIleLysAspSerGlyValAspThrCysSerSerThr	640
Db	1702	CATCGCTCTGGCAT-----	1716
Qy	641	ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys	660
Db	1717	-----AGCGATGGGCGATCCCCAATGCTCTTTGCACCTGTGACCTGGCAGCCATCCAAA	1770
Qy	661	AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal	680
Db	1771	GATGGAGATCGCCTAAATGGTGTGATATTTTAAATTAAGCGTTTAAAGATGGGAGTGA	1830
Qy	681	ProArgAspSerGlyAlaMetLeuGlyLeuLysValValGlyGlyLysMetThrGluSer	700
Db	1831	CCTCGAGATTCAGGACCAATGCTGGGCTTAAAGCTTAGGAGGAAAGATGACTGAATCA	1890
Qy	701	GlyArgLeuCysAlaPheIleThrLysValLysGlySerLeuAlaAspThrValGly	720
Db	1891	GGTCGACTTTGTGCATTTATTACCAAGTGAAGAAAGAAAGATTTAGCTGATACTGTAGGA	1950
Qy	721	HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuLeuGlnGlyAlaThr	740
Db	1951	CATCTTAGACCAAGTGTGATGAAGCTTCGGAATGGAATGGAAGCTATTGCAAGGAGCCACA	2010
Qy	741	PheGluGluValTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuVal	760
Db	2011	TTTGAGGAAGTTTACAAACATTTCTAGAATCCAAGCCTGAACACACAAGTTGAGCTTGTT	2070
Qy	761	ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu	780
Db	2071	GTFTTCAAGGCGGATTTGAGATATGCTTAGAATACCTGATGACATCATGCACAACATGGAA	2130
Qy	781	SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer	800
Db	2131	TCCAGTTCTAGCTCATTTTGAATCTCAAAAATGGATGCTCGTCTATATCGTTACTCT	2190
Qy	801	ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSerIle	820
Db	2191	CCCATGAGTCTCGCATGCTGAGGGATGTCACACAGTTCTTATCTGCACAGCTTTCAATA	2250
Qy	821	LysLeuTrpPheAspLysValGlyHisGlnLeuIleValThrIleLeuGlyAlaLysAsp	840
Db	2251	AAACTATGTTTGAACAAGTTGGTGCACAGTTATATAGTTTCAATTTTGGGAGCAAAAGAT	2310
Qy	841	LeuProSerArgGluAspGlyValGluProArgAsnProTyrValLysIleTyrPheLeuPro	860
Db	2311	CTCCCTTCCAGGGAAGATGGAGGCCAAGAAATCCTTATGTGTAATTTACTTCTCTCCA	2370
Qy	861	AspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLys	880
Db	2371	GACAGAAGTGATAAAACAAGAGAAGAAACAAAACAGTCAAGAAAACTTTGGAAACCCAAA	2430
Qy	881	TrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArgGluArgMetLeu	900
Db	2431	TGGAAACAGACTTTTCATTTATTCCTCTGTTCCACGAAGAGAAATTCGCGGAACGAATGCTC	2490
Qy	901	GluIleThrLeuTrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGlu	920
Db	2491	GAATATCACCCCTTTGGATCAAGCTCGAGTTGAGAGGAAGAAGTGAATTTCTTAGAGAG	2550
Qy	921	IleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisIleTyrLysLeuGln	940
Db	2551	ATTTTAAITGAATTTGAAAACAGCTCTGTGTAGATGATGAACACACTGTCACAACTTCAG	2610
Qy	941	ThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeuProArgArgGln	960
Db	2611	ACCATGATGTCTCTCATTTGCCACTCCCTCACCTTCTCCATATATGCGCACCGAGACAG	2670
Qy	961	LeuHisGlyGluSerProThrArgGluGlnArgSerLysArgIleSerAspSerGlu	980
Db	2671	CTGCATGGAGAGACCCAAACACGAGGTTGCAAGGTCGAAGAGAAATAGATGAACGCGAA	2730

Qy	981	ValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGly	1000
Db	2731	GTATCCGACTATGACTCGAGGACGGCGTGGAGTCTGTGCAGATTACCGACACGTGC	2790
Qy	1001	ArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCys	1020
Db	2791	CGTGATCTTCAGAGCTCCACATTATCAGTCCAGACAGCACTCATGTCATCAACCATTC	2850
Qy	1021	SerProSerGlySerProHisArgValAspValIleGlyArgThrArgSerTrpSerPro	1040
Db	2851	TCACCATCAGGCTCTCTCATCGAGTAGATGTTATAGGAAGGACTAGGTCATGTCGCCT	2910
Qy	1041	SerAlaProProProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGly	1060
Db	2911	AGTGTCCTCTCTCTCAAAGGAATGTGGAAACAGGGGCTTCGAGGGACACGTGCTACTGGC	2970
Qy	1061	HisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSer	1080
Db	2971	CATTACAAATCAATTAGCCGATGGACAGACACCGTGTCTATGGACGACCATTACTCTTCA	3030
Qy	1081	AspArgAsp	1083
Db	3031	GAGAGACAGACTCATTTTCTCACTCTACTCTGCTCTCGACACAGGCAGACCATGTGACAT	3090
Qy	1084	-----ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArg	1098
Db	3091	CACCACAGGATGGAAGGGATTGTGAAGCAGCAGATAGACAGCCATATCAAGATCCAGA	3150
Qy	1099	SerThrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArgSerSerGluArg	1118
Db	3151	TCAACAGAACACGGCTCTCTTAGAGCGGACACACCCGCTCCAGATCCTCTGACAGCT	3210
Qy	1119	ProAspThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProPro	1138
Db	3211	GCTGATACAACTCATGAGTGCAGTCTTCATTATGACTGGAGACTGCGCCCTCTCT	3270
Qy	1139	SerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThrSerProSerSer	1158
Db	3271	TCACCTGCCTTATCGAGTCTCACTCCCTCGCACTGGCTCTGTCCAGACCACTCGCTCAAGT	3330
Qy	1159	ThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeu	1178
Db	3331	ACTCCCTTAACAGGACAGAGGGGCGGCGACTCTCTCAGCTCCACCAAGGGAACACTG	3390
Qy	1179	GluArgSerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLys	1198
Db	3391	GAAGA	3396
Qy	1199	GlnValAlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGly	1218
Db	3396	-----	3396
Qy	1219	TrpAspProHisArgGlyAlaAspThrValSerThrLysSerSerAspSerAspValSer	1238
Db	3396	-----	3396
Qy	1239	AspValSerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMet	1258
Db	3396	-----	3396
Qy	1259	SerValGlnSerGluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMet	1278
Db	3396	-----	3396
Qy	1279	GlnAsnArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGly	1298
Db	3396	-----	3396
Qy	1299	AspMetCysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeu	1318
Db	3396	-----	3396
Qy	1319	GlyThrSerGlyLysLysArgSerSerIleGlyAlaLysMetValAlaIleValGly	1338

QY 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGluLysGluProGlnThrLys 100
 DB 229 -----AAG 231
 QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGluGluSerGln 120
 DB 232 CTGCATCAACAAATTTGAAATGATTAAGAGCAAGTCAAGAAATGGGAGCAATCACAA 291
 QY 121 GlnGlnGlnGlnLysGluValAspAlaProThrCysGluVileCysHisLysThrLysPhe 140
 DB 292 CAGCAGCAAGACAGAGGGCGACCCCGACCTGGCGCATCTGCCACAGCAAAATTT 351
 QY 141 AlaAspGlyCysGluHisAsnCysSerTyrCysGlnThrLysPheCysAlaLysCysGly 160
 DB 352 CGAGATGGATCGGCCCAATCTTCAATTCGCAAAACCAAGTCTTGCTGCTGTTGCGA 411
 QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
 DB 412 GGTGAGTGTCTATTACGCTCAACCAAGGTTATGTGGTGTGTAAATTTGTGCCGAAACAA 471
 QY 181 GlnGluIleLeuThrLysSerGluValaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
 DB 472 CAAGAAATCCTCACTAAGTCGGCGCGTGTGTTTCAATAGTGGGTCTTAATACCGCG 531
 QY 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGluLysLys 220
 DB 532 CAGCCTGATCAAAAGGCTCTTCAGGGCTTCGAGTGGAGAGCCCTCCAGGAGAGAG 591
 QY 221 AlaLysLeuHisGlnGlnProGlnPheGlnGlnAlaProGlyAspLeuSerValProAla 240
 DB 592 GCAAAATCGCATGAGCAGACGAGTTCAGAGGACCCCGGTGACTCATCATCTGCTGCA 651
 QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrIleLysAsnGlySerGly 260
 DB 652 GTTGAGAGAGGTCGAGCTCATGGGCTCAACAGCAGGATTCATTAAAGAAATGGCTCAGA 711
 QY 261 ValLysHisGlnIleAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
 DB 712 ATGAAGCACCAGATCCGCAATGACATGCCCTCCAGACAGAAAGAGTCCGTCACTGTC 771
 QY 281 ArgAspGlnAsnArgTyrGluGlnSerGluGluArgLysAspTyrSerGlnTyrVal 300
 DB 772 AGGGATCAAAATCGAAGATACCAACCAAGTGAAGAAAGAGAGGAATATTACAGTATGT 831
 QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSerGlnArg 320
 DB 832 CTTTCAGATAGCACATGCTAGATCTCCATCAGATTATGCTGATAGACATCTCAAGCT 891
 QY 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgArgGly 340
 DB 892 GAACCTCAATTTATGAAGACCTGATCATTTAAATTTATAGGATTTCAACAGGAGAGC 951
 QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
 DB 952 CATAGACATTTCAAAGAGTATATTGTAGACGACGAAAGTGTGGAGCAGCAGATGAATAT 1011
 QY 361 GluArgGlnArgArgGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
 DB 1012 GAAGACAAAGAGAGAGAGGAGTACCAGGACGCTACAGAAAGTATCAAAATTTGGCC 1071
 QY 381 ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer 400
 DB 1072 CGGTATCCGGTAAGCCCAACCCATATGAAGCAAAATGGGATCCACGCTGAGTGTCC 1131
 QY 401 ArgAlaArgHisGluArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
 DB 1132 CGGGCAGCAGCACAGAGAGGACAGTGTATTTCTTTGGCAAAATGCTGAATAGAT 1191
 QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
 DB 1192 TCCAGGATTTCTCTGTATAGGATGGATAGCACCATCAAGCAAGATCTGTATCTGACGT 1251
 QY 441 ArgAlaAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460

DB 1252 AGAGCTGCAATGGAAACACGAGATCGTATTCAATGGAAAGAACTCGAGAGCTCAGGA 1311
 QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
 DB 1312 CAAGATTTCTTATCCCAAGAGCACCAATCATATAGTCTCTTACCTTCGCGAGAGCCCT 1371
 QY 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLysLeu 500
 DB 1372 ATACCGCTCGATAGACAGAGCTGAGGCTGCCGACTCCCTACGGAAACAACACCACTTA 1431
 QY 501 AspProSerSerAlaValArgLysThrLysArgGluLysMetGluThrMetLeuArgAsn 520
 DB 1432 GATCCACGCTCTGCTGTAAAGAAACGAGCGAGAAAAATGGAACAATGTTGAGGAAT 1491
 QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProProProHisLys 540
 DB 1492 GATCTTTGAGTTCAGACCGCTGAGTCAGTCAGTGGAGCCGCCCCACCAAGGCCCTCATAA 1551
 QY 541 SerLysLysGlyLysMetArgGlnValSerLeuSerSerSerGluGlnGluLeuAla 560
 DB 1552 TCCAGAAAGAGAGTAAATGCGCAGGTTTCACTGACAGCTCAGAGAGAGAGTTGGCA 1611
 QY 561 SerThrProGluTyrThrSerCysAspAspValGluLeuGluSerGluSerValSerGlu 580
 DB 1612 TCCACGCTGAGTATACAACTGTGATGACGTGGAGATTGAAGCGAGAGCGTAGTGTAG 1671
 QY 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
 DB 1672 AAAGAGACATGGAG----- 1686
 QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
 DB 1687 -----TACAGCTGGTTGGAG 1701
 QY 621 AspLeuGluTrpSerGluProGlnLysAspSerGlyValAspThrCysSerSerThr 640
 DB 1702 CATGGCTCTGGCAT----- 1716
 QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
 DB 1717 -----AGCAGTGAAGGCATCCCAATGCTTTTGCAACCTGTGACCTGGCGAGCATCCAA 1770
 QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuLeuLysArgLeuLysAspGlySerVal 680
 DB 1771 GATGAGATCGCTTAATTTGGTGTATTTTAAATAAGCGTTTAAAGATGGAGGTGA 1830
 QY 681 ProArgAspSerGlyAlaMetLeuLysValValGlyGlyLysMetThrGluSer 700
 DB 1831 CCTCGAGATTCCAGGAGCAATCTGGGCTTAAAGGTTGTAGGAGGAAAGATGACTGAATCA 1890
 QY 701 GlyArgLeuCysAlaPheIleThrLysValLysLysGlySerLeuAlaAspThrValGly 720
 DB 1891 GGTGACCTTTGTGCATTTTATCCAAAGTGAAGAAAGAGAGTTTACCTGATCTAGGA 1950
 QY 721 HisLeuArgProGlyAspGluValLeuGluTrpAsnGlyArgLeuGluGlnGlyAlaThr 740
 DB 1951 CATCTTAGACAGGTCATGAGTCTTGAAATGAATGAAGCTTATTCAGAGAGCCACA 2010
 QY 741 PheGluGluValTyrAsnIleIleLeuGluSerLysProGluProGlnValGluLeuVal 760
 DB 2011 TTTGAGGAAGTTTACCAATATTCTAGATCCCAAGCTGAACCAACCAAGTTGAGCTGT 2070
 QY 761 ValSerArgProIleGlyAspIleProArgIleProAspSerThrHisAlaGlnLeuGlu 780
 DB 2071 GTTTCAGGCGGATTTGGAGATATGCTTAGAATACCTGTAGTACCTCATGSCCAATCGAA 2130
 QY 781 SerSerSerSerPheGluSerGlnLysMetAspArgProSerIleSerValThrSer 800
 DB 2131 TCCAGTTCTAGTCTATTGAATCTCAAAAAATGGATCTGCTGTATATCGGTACCTCT 2190
 QY 801 ProMetSerProGlyMetLeuArgAspValProGlnPheLeuSerGlyGlnLeuSer --- 819

Db	2191	CCCATGAGTCTCGCATGCTGAGGGATGCCACAGTTCTTATCTGGACAGCTTTCAAGC	2250
Qy	820	-----	824
Db	2251	CAAAAGCCTTAGTAAAGAACCAACGCCCTTTTGTTCCTAGGGTTCCAGATATAAATATGGTTT	2310
Qy	825	AspLysValGlyHisGlnLeuValThrIleLeuGlyAlaLysAspLeuProSerArg	844
Db	2311	GACAGGTTGGTACCCAGTTAATAGTTACAAATTTGGAGCAAGGATCTCCCTTTCAGG	2370
Qy	845	GluAspGlyArgProArgAsnProTyrValLysIleTyrPheLeuProAspArgSerAsp	864
Db	2371	GAAGATGGAGGCCAAGGAATCTTATGTTAAATTTACTTCTTCAGACAGAGTGTAT	2430
Qy	865	LysAsnLysArgArgThrLysThrValLysLysThrLeuGluProLysTrpAsnGlnThr	884
Db	2431	AAAACCAAGAGAGAACAAACACAGTCAAGAAAATTTTGGAAACCAATGGAACACAGACT	2490
Qy	885	PheIleTyrSerProValHisArgGluPheArgGluArgMetLeuGluIleThrLeu	904
Db	2491	TTCAATTTATCTCTGTTTCCCGAAGAGAAATTCGGGACGAATGCTCGAAATCACCCCTT	2550
Qy	905	TrpAspGlnAlaArgValArgGluGluSerGluPheLeuGlyGluIleLeuIleGlu	924
Db	2551	TGGGATCAAGCTCGAGTTCGAGAGGAAGATGAATTCCTAGGAGAGATTTAATTGAA	2610
Qy	925	LeuGluThrAlaLeuLeuAspAspGluProHisTrpTyrLysLeuGlnThrHisAspVal	944
Db	2611	TTGAAAACAGCTCTGTGTAGATGATGAACACACACTGGTACAAATTCAGACCCATGATGTC	2670
Qy	945	SerSerLeuProLeuProArgProSerProTyrLeuProArgGlnLeuHisGlyGlu	964
Db	2671	TCCTCATGCCACTCCCTCCACCTTCTCCATATATGCCGAGACAGCTGCATGGAGAG	2730
Qy	965	SerProThrArgArgLeuGlnArgSerLysArgIleSerAspSerGluValSerAspTyr	984
Db	2731	AGCCCAACACGACAGTTGCCAAGGTCGAAGAGATAGTGACACGGAATATCCGACTAT	2790
Qy	985	AspCysGluAspGlyValGlyValValSerAspTyrArgHisAsnGlyArgAspLeuGln	1004
Db	2791	GACTGCGAGACGGCGTGGAGTCCGTCTCAGATTACCACACGACGATGGCCGTGATCTTCAA	2850
Qy	1005	SerSerThrLeuSerValProGluGlnValMetSerSerAsnHisCysSerProSerGly	1024
Db	2851	AGCTCCACATTATCAGTGCCAGAACAAAGTCAATGTCATCAACACCAATGCTCACCATCAGGC	2910
Qy	1025	SerProHisArgValAspValIleGlyArgThrArgSerTrpSerProSerAlaProPro	1044
Db	2911	TCTCTCATCGAGTAGATGTTATAGGAAGGACTAGGTCAATGCTCGCCTAGTCCCTCT	2970
Qy	1045	ProGlnArgAsnValGluGlnGlyHisArgGlyThrArgAlaThrGlyHisTyrAsnThr	1064
Db	2971	CCTCAAAGGAATGTGGAACAGGGGCTTCGAGGGACAGCTGCTACTGGCCATTACATACA	3030
Qy	1065	IleSerArgMetAspArgHisArgValMetAspAspHisTyrSerSerAspArgAsp---	1083
Db	3031	ATTAGCCGATGGACAGACACCGTGTATGGACGACCAATTAATCTTTCAGAGAGAGACAGT	3090
Qy	1083	-----	1083
Db	3091	CATTTTCTCACTTCACTCGCTCTCGACACAGGCAGACCATGAGCATCACACAGGGAT	3150
Qy	1084	---ArgAspCysGluAlaAlaAspArgGlnProTyrHisArgSerArgSerThrGluGln	1102
Db	3151	GGAAGGGATTGTGAAGCAGCAGATAGACAGCCATATCACAGATCCAGATCAACAGAACAA	3210
Qy	1103	ArgProLeuLeuGluArgThrThrArgSerArgSerSerGluArgProAspThrAsn	1122
Db	3211	CGGGCTCTCTTAGACGGGACCAACACCGCTCCAGATCCTCTGAAACGTCGTGATACAAAC	3270
Qy	1123	LeuMetArgSerMetProSerLeuMetThrGlyArgSerAlaProProSerProAlaLeu	1142
Db	3271	CTCATGAGGTGATGCCCTTCAATTAATGACTGGAGAGATCTGCCCTCTTCACTGCCCTTA	3330
Qy	1143	SerArgSerHisProArgThrGlySerValGlnThrSerProSerSerThrProGlyThr	1162
Db	3331	TCAGGTCTCACCTCGCACTGTGCTGTCCAGACAGTCCGTCAGTACTCCCGTAACA	3390
Qy	1163	GlyArgArgGlyArgGlnLeuProGlnLeuProProLysGlyThrLeuGluArgSerAla	1182
Db	3391	GGACGAAGGGCGCGGAGCTTCTCTCAGCTCCCAACAAAGGACACTGGAAAGA	3444
Qy	1183	MetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnValAlaGly	1202
Db	3444	-----	3444
Qy	1203	SerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTrpAspProHis	1222
Db	3444	-----	3444
Qy	1223	ArgGlyAlaAspThrValSerThrLysSerSerAspSerValSerAspValSerAla	1242
Db	3444	-----	3444
Qy	1243	ValSerArgThrSerSerAlaSerArgPheSerSerThrSerTyrMetSerValGlnSer	1262
Db	3444	-----	3444
Qy	1263	GluArgProArgGlyAsnArgLysIleSerValPheThrSerLysMetGlnAsnArgGln	1282
Db	3444	-----	3444
Qy	1283	MetGlyValSerGlyLysAsnLeuThrLysSerThrSerIleSerGlyAspMetCysSer	1302
Db	3444	-----	3444
Qy	1303	LeuGluLysAsnAspGlySerGlnSerAspThrAlaValGlyAlaLeuGlyThrSerGly	1322
Db	3444	-----	3444
Qy	1323	LysLysArgSerSerIleGlyAlaLysMetValAlaIleValGlyLeuSerArgLys	1342
Db	3444	-----	3444
Qy	1343	SerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLysLeuArgSerThr	1362
Db	3445	-----GGAGGAGTAAAGCTACGACGACT	3471
Qy	1363	ValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThrArgGln	1382
Db	3472	GTCCAGAGAAGCACCGAGACCCGGCTAGCAGTGGAGATGAGGAACCTGGATGACAAGACAG	3531
Qy	1383	AlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsnLeuIle	1402
Db	3532	GCCAGCCGGGAATCCACGGATGCCAGATGAAACAGTTACAGCTCAGAAGGAAATCTGATC	3591
Qy	1403	PheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGlyLeuGly	1422
Db	3592	TTTCTCGGGCTCGCCCTGCGCTCTGACAGCCAGTTTCTTGAATGCTTGGGC	3651
Qy	1423	ProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIleGlnVal	1442
Db	3652	CCTGCTCAGCTAGTGGGACGCCAGCCCTGGGACTCTCTGCAATGGGTGACATTCAGGTG	3711
Qy	1443	GlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAlaArgGlyLeuVal	1462
Db	3712	GGGATGATGATGAAGAGGACAGCTGGAGGTAGAGATCATCCGGCTCGAGCCCTTGTA	3771
Qy	1463	ValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeuAspAsn	1482
Db	3772	GTAACACCGGATTCCTAAGACACTGCCAGCACCATATGTCAAGGTGACTCTGTAGACAAC	3831
Qy	1483	GlyValCysIleAlaLysLysThrLysValAlaArgLysThrLeuGluProLeuTyr	1502
Db	3832	GGAGTCTGCATAGCCAAAGAAACCAAGGTGGCGAAGAGACGCTGGAGCCCTGTAT	3891

Qy 1503 GlnGlnLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleValTrp 1522
 Db 3892 CAGCAGCTGTTATCTTTTCAAGAGAGAGCCCGAGGAGAGGTTTACACATCATTTGCTCG 3951
 Qy 1523 GlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeuAsp 1542
 Db 3952 GGAGATTATGCGCGCATGATGACAAATCCCTTTATGGAGTGGCCCATCTCTTAGAT 4011
 Qy 1543 GluLeuGluLeuSerAspMetValIleGlyTrpPheLysLeuPheProPheSerSerLeu 1562
 Db 4012 GAATGGAACTATCCAAATGTTATGATGGTTCAAACTCTTCCCTCTCTCCCTCCTA 4071
 Qy 1563 ValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGluSerSer 1582
 Db 4072 GTAGATCCAACTTGGACCTCTGACAAAGAGAGCTCCCAATGCTCTCTGAAAGTTCT 4131
 Qy 1583 ThrGlyProSerTyrSerArgSer 1590
 Db 4132 ACTGACCTTCTACTCTCGTTCA 4155
 RESULT 14
 AF548738 5063 bp mRNA linear ROD 12-MAR-2003
 LOCUS Rattus norvegicus synaptic regulatory protein RIM2beta (Rim2beta)
 DEFINITION mRNA, complete cds.
 ACCESSION AF548738
 VERSION AF548738.1 GI:24415909
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 5063)
 Wang, Y. and Sudhof, T.C.
 TITLE Genomic definition of RIM proteins: evolutionary amplification of a
 family of synaptic regulatory proteins
 JOURNAL Genomics 81 (2), 126-137 (2003)
 MEDLINE 22508184
 PUBMED 12620390
 2 (bases 1 to 5063)
 Wang, Y. and Sudhof, T.C.
 TITLE Direct Submission
 JOURNAL Submitted (30-SEP-2002) Center for Basic Neuroscience, The
 University of Texas Southwestern Medical Center at Dallas, 6000
 Harry Hines Blvd, Dallas, TX 75390, USA
 FEATURES
 location/Qualifiers
 1..5063
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 1..5063
 /gene="Rim2beta"
 51..4142
 /gene="Rim2beta"
 /codon_start=1
 /product="synaptic regulatory protein RIM2beta"
 /protein_id="AA55930.1"
 /db_xref="GI:24415910"
 /translation="MQFETLRQVNCVSLHFHGFVSPNPNTLQNELFQTLNNARKS
 PSVRDQNRDYQSEEEYQYVPSDSTMPSPSDYADRSQREPOFYEPDFHLNTR
 DSNRGRHRSKEYIVDDSDVSDEYERQREERYQARYSDPNLARVPVKPYEBO
 MRHAIEVSRAHERHSDVSLANAELSDRIILLRMPSPRQSVSEERAAWENORSY
 SMERTREAGSSVPQRTNPSPTPRSPRPLDPELRADSLAKQHLLDPSAVRK
 TKREHMETLNLNLSQSSSEVRPPPRPHKSKGKQKQVSLSSSEELASTPEYT
 SCDDVIESSEVGEKMEYSLWEHSHSSEASPMSLHPVTWOPSKGDLRIGILL
 NKRLLKDSVPDSDGMLGLVVGKWTSGRLCAFTKVKKGLADTVGHLRPGDEVL
 EWNGRLQAGTFFVYNTIILSKPEPQVELVVRPIGMPIRIPDSTHAQLSSSSSFE
 SQWDRPSIVTSQWSPGMLRDVPOFLSGQLSSLSRRTPTFPVFOVKLWPKVGH
 QLVITLGAKLDPSPDKRPYVYKIFLPSRDKNKRRTKTKVKTLEPKWNTQFIY
 SPVHRREFRPMRLQLHGESPTRLQRSKRIISDEVSVDCEGDGVGVVSYRHPGRDL
 SLPLPHSPSPRQLHGESPTRLQRSKRIISDEVSVDCEGDGVGVVSYRHPGRDL
 QSSTLSVPEQVMSNNHCSFGSPGPHRVDVIGTRTWSVSPVPPQRNVEQLGRTATGCH

YNTISRMDRHRVMDHYSSERDSHFLTPRSRHRQTSSEHHRDGRCEAADROPYHRS
 RSTQRELLERTTTRSSRSADTNMRSPSLMTGSRASAPPSPALSRHPTQSVQTS
 PSTPTVTRGRROLPOLPKGTGLERMITTDMDSSTRKNSGAMDIETERNRQMKLNKYKO
 VAGSDPELEODVHSKYRSCWDPHRGADTVSTKSDSDPSVSAVERTSSASRFSSTSY
 MSVQSERPRGNRLKISVFTSKQSRQMGVSGSKMAKSTSISGDMCSLEKNDGSSDITAV
 GALTGSKKRSISIGAKWVAIVGLSRKSRASQLSQTESGGKKLSTVQSTETGLAV
 EMNMTWTRQSRSTTSGMNSYSBGNLIIFPVRLASDSQSFDFLGLGPAQLVGRQT
 LATPAMGDIQVGMWDKKGQLEBVIIRAGLVKVPKSKTLTPAPYKVVLLDNGVCIACK
 KTKVARKTLPLQYQQLLSPEESQGVQLQIIVMGDYGMDHDKFPMGVQVAILLDELELS
 NMVIGWPKLFPSPSLVDPTLAPITRRASQSSLSSTGSPSISRS"
 BASE COUNT 1547 a 1130 c 1238 g 1148 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.19e-188 Length: 5063
 Score: 6261.50 Matches: 1230
 Percent Similarity: 91.02% Conservative: 17
 Best Local Similarity: 89.78% Mismatches: 24
 Query Match: 75.58% Indels: 99
 DB: 1.0 Gaps: 5
 US-09-617-099B-1 (1-1590) x AF548738 (1-5063)
 Qy 273 ArgLysArgSerProSerValSerArgAspGlnAsnArgArgTyrGluGlnSerGluGlu 292
 Db 171 AGAAAAAGAAATCGTCCAGTGTCCAGGATCAAAATCGAAGATACGACCAAGTGAAGAA 230
 Qy 293 ArgGluAspTyrSerGlnTyrValProSerAspGlyThrMetProArgSerProSerAsp 312
 Db 231 AGAGAGAATATTCACAGTATGTTCTTCAGATAGCAATGCTCTAGATCTCCATCATGAT 290
 Qy 313 TyrAlaAspArgArgSerGlnArgGluProGlnPheTyrGluGluProGlyHisLeuAsn 332
 Db 291 TATGCTGATAGACCATCTCAACGTGAACCTCAATTTATGAGAACTGATCATTTAAT 350
 Qy 333 TyrArgAspSerAsnArgArgGlyHisArgHisSerLysGluTyrIleValAspAspGlu 352
 Db 351 TATAGGATTTCTAAGAGAGAGGCGCATAGACATTTCCAAAGAGATATTTGTAGACGCGAA 410
 Qy 353 AspValGluSerArgAspGluTyrGluArgGlnArgGluGluGluTyrGlnAlaArg 372
 Db 411 GATGTGAGAGACAGAGATGAATATGAAAGACAAAGGAGAGAGAGAGTACCGGACGC 470
 Qy 373 TyrArgSerAspProAsnLeuAlaArgTyrProValLysProGlnProTyrGluGluGln 392
 Db 471 TACAGAGATGATCCAAATTTGGCCCGGTATCCGGTAAAGCCACCACTATGAGAGAGAA 530
 Qy 393 MetArgIleHisAlaGluValSerArgAlaArgHisGluArgArgHisSerAspValSer 412
 Db 531 ATCGGATCCACGCTGAAGTGTCCCGGCGACGACACGAGAGAGGACAGTGTGTTCT 590
 Qy 413 LeuAlaAsnAlaGluLeuGluAspSerArgIleSerLeuLeuArgMetAspArgProSer 432
 Db 591 TTGGCAAATGCTGAAGATAGAGATTTCCAGGATTTCTCTGCTAGAGATGATAGACCATCA 650
 Qy 433 ArgGlnArgSerValSerGluArgArgAlaAlaMetGluAsnGlnArgSerTyrSerMet 452
 Db 651 AGGCAAGATCTGTATCTGAACGTAGAGCTGCAATGGAAACCAAGCATGCTATTCAATG 710
 Qy 453 GluArgThrArgGluAlaGlnGlnSerSerTyrProGlnArgThrSerAsnHisSer 472
 Db 711 GAAAGAACTCGAAGAGCTCAGGAGCAAAAGTTCTTTTATCCACAAAGACCAACATCATAGT 770
 Qy 473 ProProThrProArgArgSerProIleProLeuAspArgProAspMetArgArgAlaAsp 492
 Db 771 CCTCTTCCCTCGCAGAGAGCCCTATACCGCTCGATAGACCAAGAGCTGACGCGTCCGAC 830
 Qy 493 SerLeuArgLysGlnHisHisLeuAspProSerSerAlaValArgLysThrLysArgGlu 512
 Db 831 TCCCTACGGAACCAACACCATCTAGATCCCAAGCTCTCTGCTGTAAGAAACCAAGACGAGAA 890
 Qy 513 LysMetGluThrMetLeuArgAsnAspSerLeuSerSerAspGlnSerGluSerValArg 532

891 AAAATGGAAACAATGTTGAGGAATGATCTTTGAGTTTCAGACCAGTCTGAGTCAGTGAGG 950
533 ProProProArgProHisLysSerLysGlyLysMetArgGlnValSerLeu 552
951 CCGCCCCCAAGGCTCATTAATCAAGAGGAGTAATGCGCCAGGTTTCACTG 1010
553 SerSerSerGluGluGluLeuAlaSerThrProGluTyrThrSerCysAspValGlu 572
1011 AGCAGCTCAGAGGAGGTTGGCATCCAGCCTGAGTATACAAGCTGTGATGACGTGGAG 1070
573 LeuGluSerGluSerValSerGluValGlyValSerGlnLysGlyLysArgLysThrSer 592
1071 ATTGAAGCAGAGCGTAGGTGAGAAAGGAGCATGCGG----- 1109
593 GluGlnGlyValLeuSerAspSerAsnThrArgSerGluArgGlnLysLysArgMetTyr 612
1109 ----- 1109
613 TyrGlyGlyHisSerLeuGluGluAspLeuTyrTrpSerGluProGlnLysLysAspSer 632
1110 -----TACAGCTGGTTGGAGCATGCGCTTTGGCAT----- 1139
633 GlyValAspThrCysSerSerThrThrLeuAsnGluGluHisSerHisSerAspLysHis 652
1140 -----AGCAGTGAGGCATCCCCAATGCTTTGCAC 1169
653 ProValThrTrpGlnProSerLysAspGlyAspArgLeuLeuGlyArgLysLeuLeuAsn 672
1170 CCGTGACCTGGCAGCATCCAAAGATGGAGATCGCTTAATTGGTGTGATTTTATTAAT 1229
673 LysArgLeuLysAspGlySerValProArgAspSerGlyAlaMetLeuGlyLysVal 692
1230 AAGCGTTTAAAGATGGGAGTGTTACCTCGAGATTTCAGGAGCAATGCTGGGCTTAAAGGTT 1289
693 ValGlyLysMetThrGluSerGlyArgLeuCysAlaPheIleThrLysValLysLys 712
1290 GTAGGAGAAAGATGACTGAATCAGGTGCGACTTTGTGCATTTTATACCAAGTGAAAGAA 1349
713 GlySerLeuAlaAspThrValGlyHisLeuArgProGlyAspGluValLeuGluTyrAsn 732
1350 GGAAGTTTAGCTGATCTGTAGGACATCTTAGACCGAGTGATGAAGCTTGGATGGAT 1409
733 GlyArgLeuLeuGlnGlyAlaThrPheGluGluValTyrAsnLysLysLeuGluSerLys 752
1410 GGAAGGCTATTGCAAGGAGCCACATTTGAGGAGTTTACACATTAATTTAGAAATCCAAAG 1469
753 ProGluProGlnValGluLeuValValSerArgProIleGlyAspIleProArgIlePro 772
1470 CCTGAACCAAGTTGAGCTTGTGTTTCAAGGCGGATGGAGATATGCTTAGAATACCT 1529
773 AspSerThrHisAlaGlnLeuGluSerSerSerSerPheGluSerGlnLysMetAsp 792
1530 GATAGCACTCATGCAACCTGGAATCCAGTTCTAGCTCATTTGAATCTCAAAAATGGAT 1589
793 ArgProSerIleSerValThrSerProMetSerProGlyMetLeuArgAspValProGln 812
1590 CGTCCGCTATATCGGTTACCTCTCCCATGAGTCTCGCATGCTGAGGAGTGTCCACAG 1649
813 PheLeuSerGlyGlnLeuSer----- 819
1650 TTTCTATCTGACACGCTTTCAAGCCAAAGCCTTAGTAGAAGAACAAAGCCTTTTGTCTCT 1709
820 -----IleLysLeuTyrPheAspLysValGlyHisGlnLeuIleValThrIleLeu 836
1710 AGGGTTTCAGATAAACTATGTTTGACAAAGTTGCTCACCAGTTAATAGTTACAATTTTG 1769
837 GlyAlaLysAspLeuProSerArgGluAspGlyArgProArgAsnProTyrValLysIle 856
1770 GGAGCAAGAGTCTCCCTTCCAGGAGAGATGGAGGCCCAAGGAATCTTATGTTAAATTT 1829
857 TyrPheLeuProAspArgSerAspLysAsnLysArgArgThrLysThrValLysLysThr 876
1830 TACTTTCTTCCAGACAGAGTGTATAAAAACAAGAGAGAAACAAACAGTCAAGAAAACT 1889

877 LeuGluProLysTrpAsnGlnThrPheIleTyrSerProValHisArgArgGluPheArg 896
1890 TTGGAAACCAAAATGGAAACAGACTTTTCAATTATTTCTCTCTTCCAGGAAGAAATTCGCG 1949
897 GluArgMetLeuGluIleThrLeuTyrAspGlnAlaArgValArgGluGluSerGlu 916
1950 GAACGAATGCTCGAATCACCTTTGGATCAAGCTCGAGTTCGAGGAGAAAGTGAA 2009
917 PheLeuGlyGluIleLeuIleGluLeuGluThrAlaLeuLeuAspAspGluProHisTyr 936
2010 TTCTTAGGAGAGATTTTAATTGAATTGGAACAGCTCTGTAGATGATGAACCACTGG 2069
937 TyrLysLeuGlnThrHisAspValSerSerLeuProLeuProArgProSerProTyrLeu 956
2070 TACAAACTTCAGACCCTCATGTCCTCATTTGCCACTCCCTCACCTTCTCCATATATG 2129
957 ProArgArgGlnLeuHisGlyGluSerProThrArgArgLeuGlnArgSerLysArgLys 976
2130 CCACGGAGACAGCTGCATGGAGAGACCCCAACAGCAGGTTGCAAGGTCGAAGAGATA 2189
977 SerAspSerGluValSerAspTyrAspCysGluAspGlyValGlyValValSerAspTyr 996
2190 AGTGACAGCGAAGTATCCGACTATGACTCGAGGACGCGTGGAGTCTGTGCAGATTAC 2249
997 ArgHisAsnGlyArgAspLeuGlnSerSerThrLeuSerValProGluGlnValMetSer 1016
2250 CGACAGATGGCGCTGATCTTCAAGCTCCACATTAATCATGTGCAGAACAAAGTCATGTCA 2309
1017 SerAsnHisCysSerProSerGlySerProHisArgValAspValIleGlyArgThrArg 1036
2310 TCAACCATTTGCTCACCCTCAGCTCTCTCTCATCGAGTAGATGTTTATAGGAGACTAGG 2369
1037 SerTrpSerProSerAlaProProGlnArgAsnValGluGlnGlyHisArgGlyThr 1056
2370 TCATGTGCGCTAGTGTCCCTCTCTCAAGGAATGTGGAACAGGCGCTTCAGGGACA 2429
1057 ArgAlaThrGlyHisTyrAsnThrIleSerArgMetAspArgHisArgValMetAspAsp 1076
2430 CGTGTACTGGCCATTACAATACAAATAGCCGATGACAGACACCGTGTCTGACGACAC 2489
1077 HisTyrSerSerAspArgAsp----- 1083
2490 CATTTACTCTTCAGAGAGAGACAGTCATTTTCTCACTCTACCTCGCTCTCGACACAGCAG 2549
1084 -----ArgAspCysGluAlaAlaAspArgGlnProTyr 1094
2550 ACCAGTGAGCATCACCAAGGATGGAGGATTTGTGAAGCAGCAGATAGACAGCCATAT 2609
1095 HisArgSerArgSerThrGluGlnArgProLeuLeuGluArgThrThrThrArgSerArg 1114
2610 CACAGATCCAGATCAACAGAACACCGCTCTCTCTAGAGCGGACCAACCCCGCTCCAGA 2669
1115 SerSerGluArgProSerThrAsnLeuMetArgSerMetProSerLeuMetThrGlyArg 1134
2670 TCCTCTGAACGTGCTGATACAAACCTCATGAGGTGATGCTCTCATTAATGACTGGAAGA 2729
1135 SerAlaProProSerProAlaLeuSerArgSerHisProArgThrGlySerValGlnThr 1154
2730 TCTGCCCCCTCTTCACTTGCCTTATCGAGTCTCACCTCGACTGCTGTGTCCAGACC 2789
1155 SerProSerSerThrProGlyThrGlyArgArgGlyArgGlnLeuProGlnLeuProPro 1174
2790 AGTCCGTCGAAGTACTCCCGTAAACAGGACGAAGGCGCGGAGCTTCTCTAGCTCCCAACA 2849
1175 LysGlyThrLeuGluArg----- 1180
2850 AAGGGAACTCTGGAAAGAAATGATCACAGAGACATGCACTCCCAAGAGAGAGAACTCA 2909
1181 SerAlaMetAspIleGluGluArgAsnArgGlnMetLysLeuAsnLysTyrLysGlnVal 1200
2910 GGTCTATGATATAGAGGAGAAATCGCCAAATGAAATTTAACAAATTAACAAACAGGTA 2969

1201 AlaGlySerAspProArgLeuGluGlnAspTyrHisSerLysTyrArgSerGlyTyrAsp 1220
 2970 GCGGATCAGATCCAGACTGAGAACAGATACCATTCGAGATGCTCAGATGGGAT 3029
 1221 ProHisArgGlyAlaAspThrValSerThrLysSerAspSerAspValSerAspVal 1240
 3030 CCACATAGAGGGCGAGATCTGTTCTACTAAATCCTCGACAGATGATGAAGTATGTA 3089
 1241 SerAlaValSerArgThrSerSerAlaSerArgPheSerSerThrSerThrMetSerVal 1260
 3090 TCTCGGTTTCAAGACTAGTAGTCTCTCTGTTTTCAGCAGACACAGCTACATGTCGTC 3149
 1261 GlnSerGluArgProArgGlyAsnArgLysLysLeuSerValPheThrSerLysMetGlnAsn 1280
 3150 CAATCAGACCGCGGAGAGAACAGCAAAATCAGTGCTTACGTCGCAAAATGCAAGC 3209
 1281 ArgGlnMetGlyValSerGlyLysAsnLeuThrLysSerThrSerLysSerGlyAspMet 1300
 3210 AGACAGATGGGTGTGCGGGGAGAGCATGGCCAAAAGCAACAGCATCAGTCGAGCATG 3269
 1301 CysSerLeuGluLysAsnAspGlySerGlnSerAspThrAlaValAlaLeuGlyThr 1320
 3270 TCTCTCAGTGGAGAAATGATGCGACCGACCTGCGACCTGCGGCGCCCTGGCGACC 3329
 1321 SerGlyLysLysArgSerSerLysLeuAlaLysMetValAlaLysValGlyLeuSer 1340
 3330 AGTGGACAGACGGCGATCTAGCATTTGGGGCCAAATGATAGCTATTGTTGCTCTCA 3389
 1341 ArgLysSerArgSerAlaSerGlnLeuSerGlnThrGluGlyGlyLysLysLeuArg 1360
 3390 CGGAAAAGTCCAGTGCCTCCCACTCAGCCAAACCGAAGGAGGAGTAAAGCTACGC 3449
 1361 SerThrValGlnArgSerThrGluThrGlyLeuAlaValGluMetArgAsnTrpMetThr 1380
 3450 AGCACTGTCCAGAGAAGCAGCGAGACCGGGCTAGCAGTGGAGATGAGGAACCTGATCA 3509
 1381 ArgGlnAlaSerArgGluSerThrAspGlySerMetAsnSerTyrSerSerGluGlyAsn 1400
 3510 AGACAGCGCCAGCGGAAATCCAGGATGGCAGATGAACAGTTACAGCTCAGAGGAAT 3569
 1401 LeuIlePheProGlyValArgLeuAlaSerAspSerGlnPheSerAspPheLeuAspGly 1420
 3570 CTGATCTTCTGGGGTCCGCTGGCTCTGACAGCCAGTTCCAGCATTTCTTGATGAT 3629
 1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440
 3630 CTGGGCGCTGCTCAGCTAGTGGAGCGCCAGACCTGGCGACTCTCTCAATGGGTGACAT 3689
 1441 GlnValGlyMetMetAspLysLysGlyGlnLeuGluValGluIleIleArgAlaArgGly 1460
 3690 CAGGTGGGATGATGATGATAGAGGAGCAGCTGAGGTAGAGATCATCCGGGCTCGAGGC 3749
 1461 LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu 1480
 3750 CTGTAGTAAACAGGTTCCAGACACTGCCAGCACCATATATGTCAGGTGTACCTGTTA 3809
 1481 AspAsnGlyValCysIleAlaLysLysLysThrLysValAlaArgLysThrLeuGluPro 1500
 3810 GACAAACGAGTCTGATAGCCAAAGAAACCAAGTGGCGAGAGAGAGCGTGGAGCGCC 3869
 1501 LeuTyrGlnGlnLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleIle 1520
 3870 CTGTATCAGCAGCTGTATCTTTCGAGAGAGAGCCCGAGGAGGTGTGTACAGATCAT 3929
 1521 ValTyrGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu 1540
 3930 GTCTGGGAGATTTATGGCGCATGATCATCAATCTTTATGGAGTGGCCGAGATCTC 3989
 1541 LeuAspGluLeuGluSerAsnMetValIleGlyTyrPheLysLeuPheProSer 1560
 3990 TTAGATGAACAGTGAACACTCAACATGATGATGATGATGATGATGATGATGATGATG 4049
 1561 SerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGlu 1580

Db 4050 TCCCTAGTAGATCAACCTTGGCACTCTGACAGAGAGCTTCCCAATGCTCTCGAA 4109
 Qy 1581 SerSerThrGlyProSerTyrSerArgSer 1590
 Db 4110 AGTTCTACTGACCTTCTTACTCTCGTTCA 4139
 RESULT 15
 AF199328 AF199328 4851 bp mRNA linear ROD 04-JUL-2000
 LOCUS Rattus norvegicus RIM2-4B (Rim2-4B) mRNA, complete cds.
 DEFINITION AF199328
 ACCESSION AF199328
 VERSION AF199328.1 GI:8925869
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 4851)
 AUTHORS Wang, Y., Sugita, S. and Sudhof, T. C.
 TITLE The RIM/NIM family of neuronal C2 domain proteins. Interactions
 with Rab3 and a new class of Src homology 3 domain proteins
 JOURNAL J. Biol. Chem. 275 (26), 20033-20044 (2000)
 MEDLINE 20347919
 PUBMED 10748113
 REFERENCE 2 (bases 1 to 4851)
 AUTHORS Wang, Y. and Sudhof, T. C.
 TITLE Direct Submission
 JOURNAL Submitted (27-OCT-1999) Center for Basic Neuroscience, The
 University of Texas Southwestern Medical Center, 6000 Harry Hines
 Blvd., Dallas, TX 75235-9111, USA
 FEATURES
 Location/Qualifiers
 1..4851
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 1..4851
 /gene="Rim2-4B"
 52..3930
 /gene="Rim2-4B"
 /codon_start=1
 /product="RIM2-4B"
 /protein_id="AAF81650.1"
 /db_xref="GI:8925870"
 /translation="MSAPLGRGRPAPTAAQPPQPPMPDLHLTLBEERKIIQAVM
 DRQKEEKEQSVLKKHQFQYKQVKGESQOQBOGKADPTGICHTKFPAD
 CGHNCYCTKFCARCGSVLRKSNKVMVNCNLCKQOEILTKGAWFNSGNTPO
 QPDOKALASERAPQEKAKLHQTQFQPGDSVPAVERGAHGLTRDSTKNG
 SGMKQIADMPDSKRPSPVSDONRRYQDSEEREYSQYPSSTWPSPSYADR
 RSQREFQYERPDHLYNRDSNRGRHSKTYIVDDVEDSEDEYERQREERYQAIR
 SDPNLARYPKPQYEEQMRTHAEVSRAPHERHSDVSLANLEDSRISLMDRPS
 RQRSVSRRAAMENQSYSMERTREAOQSSYPQRTTNHSPPTPERSPIPLDRPELR
 ADSLRKHLPDSSAVRTRKREKMTLRNDLSLSDQSESVPPPPRPHKSKGQMR
 QVLSLSSSEELASTPEYTSQDVEISESVGKEGMEYSLEHSAHSSSEAFMSLHP
 VTWQSKDGLRIGILLNKLKDGSVPRDSGLMLGLVVGEMTESGELCAPIYKVK
 KGLSATVGLHRLGDELEWNGLLQGAIFSEVINILISEKPEPVELVPSPIGMP
 RIPSTHAQLSSSSSSSQKMDPSISVTSPMSGLMDRVPQFLSGSLKLMFDKY
 GHQLVITLIGALDPSREDGRPNPYVKIYFLDRSDKNRRTKTKVTKLEPNQTF
 IYSPVRRERFRERLEITLMDQARVREESFELGELILELALDDPHYKLTQTH
 VSSLPLHPSYPMRQRLHGSPTERLQSRKISDSSEVSDYDCEDGVSVSDYRDR
 DLQSTLSVPEQVSSNCSPSGSPHYVDYIGRTNSPSPSPPPPPNVEQGLRGTAT
 GHYNTISMDRHRVMDHYSSEDSHFUTLPSRHRSTVQRETEGLAVENMWTQASRE
 PSSTVTRGRQLPQPKTLERGQKKLRSTVQRETEGLAVENMWTQASRE
 STDGSMNSYSGENLIPFQVRLASQSFDFLGLGPAQLVGRQTLATPAMDIOQGM
 MDKKQLBVEIIRAGLVVKGSKTLPAVYKVVLLDNGVCIKKTKTKVAKTLPELY
 QQLLSPSPQGVLIQIIVWDYGBMDHKSFMGVAQIILLDELSLNNWLGWTKLPPPS
 SLDFTLAPLTRASQSSLESSTGSPSISRS"
 BASE COUNT 1453 a 1110 c 1209 g 1079 t
 ORIGIN
 Alignment Scores: 4.11e-187 Length: 4851
 Pred. No.:

Score: 6212.00 Matches: 1226
Percent Similarity: 78.43% Conservative: 21
Best Local Similarity: 77.11% Mismatches: 45
Query Match: 74.98% Indels: 298
DB: 10 Gaps: 5

US-09-617-099B-1 (1-1590) x AF199328 (1-4851)

QY 1 MetSerAlaProLeuGlyProArgProAlaProThrProAlaSerGlnPro 20
Db 52 ATGTGGCTCCACTCGGGCCCGGGCGCGCCGCCCTCCAGCCT 111
QY 21 ProProGlnProGluMetProAspLeuSerHisLeuThrGluGluArgLysLeu 40
Db 112 CCTCCGAGCCCGAGATCGCGACCTCAGCCACTCAGGAGAGGAGGAAATCATC 171
QY 41 LeuAlaValMetAspArgGlnLysLeuGluGluGluGluGluGluGlu 60
Db 172 CAGGCTGTCTATGATCGTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 228
QY 61 LysGluGluHisLysAlaGlnProThrGlnTrpPheProPheSerGlyLeuThrGluLeu 80
Db 228 ----- 228
QY 81 ValAsnAsnValLeuGlnProGlnGlnLysGlnProAsnGlnLysGluProGlnThrLys 100
Db 229 -----AAG 231
QY 101 LeuHisGlnGlnPheGluMetTyrLysGluGlnValLysLysMetGlyGluGluSerGln 120
Db 232 CTGCATCAACAATTTGAATGTATPAGAGCAAGTCAAGAAATCGGAGAGGATCAAA 291
QY 121 GlnGlnGlnGlnLysGlyAspAlaProThrCysGlyLeuCysHisLysThrLysPhe 140
Db 292 CAGCAGCAAGAGCAGAGAGCGCGCCCGACCTCGCGCATCTGCCCAACAAGCAAAATTT 351
QY 141 AlaAspGlyCysGlyHisAsnCysSerTyrCysGlnThrLysPheCysAlaArgCysGly 160
Db 352 GCAGATGGATCGCGCCATAACTGTTCATATTGCCAAACCAAGTTCTGTCTCGTGTGGA 411
QY 161 GlyArgValSerLeuArgSerAsnLysValMetTrpValCysAsnLeuCysArgLysGln 180
Db 412 GGTGAGTGTCTATTACGCTCAAAACAGGTATGTGGTGTGTATTTGTGCCGAAACAA 471
QY 181 GlnGluLeuThrLysSerGlyAlaTrpPheTyrAsnSerGlySerAsnThrLeuGln 200
Db 472 CAAGAATCCTCACTAAGTCGGCGCGTGTGTTCATATAGTGGTCTATATACACCCGAG 531
QY 201 GlnProAspGlnLysValProArgGlyLeuArgAsnGluGluAlaProGlnGluLysLys 220
Db 532 CAGCCTGATCAAAAGGCTCTTCGAGGGCTTCGACGTGAGGAGAGCCCTCAGGAGAGAAG 591
QY 221 AlaLysLeuHisGlnGlnProGlnPheGlnGlyAlaProGlyAspLeuSerValProAla 240
Db 592 GCAAACTGCATGAGACAGCCAGTTCAGGAGACCCCGGTGACTCATAGTACTTGCA 651
QY 241 ValGluLysGlyArgAlaHisGlyLeuThrArgGlnAspThrLysAsnGlySerGly 260
Db 652 GTTCAGAGAGTTCGAGCTCATGGCTCACAAGACAGGATTTCTATTAGAATGGCTCAGGA 711
QY 261 ValLysHisGlnLysAlaSerAspMetProSerAspArgLysArgSerProSerValSer 280
Db 712 ATGAAGCACCCAGATCGCCAGTGCATGCTTCAGACAGAAAGAGAGAGAGAGAGAGAG 771
QY 281 ArgAspGlnAsnArgArgTyrGluGlnSerGluGluArgGluAspTyrSerGlnTyrVal 300
Db 772 AGGGATCAAAATCGAAGATACACCAAGTGAAGAAAGAGAGAGATATTCACGATATGTT 831
QY 301 ProSerAspGlyThrMetProArgSerProSerAspTyrAlaAspArgSerGlnArg 320
Db 832 CCTTCAGATAGACACAAATGCCATAGATCTCCATCAGATTATGCTGATAGACGATCTCAAGT 891
QY 321 GluProGlnPheTyrGluGluProGlyHisLeuAsnTyrArgAspSerAsnArgGly 340

Db 892 GAACTCAATTTATGAGAGACTGATCATTTAAATATTAGGATTTCTAACAGGAGGC 951
QY 341 HisArgHisSerLysGluTyrIleValAspAspGluAspValGluSerArgAspGluTyr 360
Db 952 CATAGACATTTCCAAAGAGTATATTGTAGACGACGAGAGATGTGAGAGCGAGATGAAT 1011
QY 361 GluArgGlnArgArgGluGluGluTyrGlnAlaArgTyrArgSerAspProAsnLeuAla 380
Db 1012 GAAGACAAGGAGAGAGAGAGAGTACCAAGGACGCTCAGAGATGATCCAAATTTGGCC 1071
QY 381 ArgTyrProValLysProGlnProTyrGluGluGlnMetArgIleHisAlaGluValSer 400
Db 1072 CUGTATCCGTTAAGCCACACCTTATGAAGCAATTCGGATCCACGCTGAGTGTTC 1131
QY 401 ArgAlaArgHisGluArgHisSerAspValSerLeuAlaAsnAlaGluLeuGluAsp 420
Db 1132 CGGCGACGACACGAGAGAGGACAGTGTATTTCTTGGCAATGTCTGAATCTAGAAT 1191
QY 421 SerArgIleSerLeuLeuArgMetAspArgProSerArgGlnArgSerValSerGluArg 440
Db 1192 TCCAGGATTTCTCTGCTAAGGATGGATAGACCATCAAGGCAAGATCTGTATCTGAACGT 1251
QY 441 ArgAlaMetGluAsnGlnArgSerTyrSerMetGluArgThrArgGluAlaGlnGly 460
Db 1252 AGAGCTGCAATGGAACCCAGCGATCTGTATTCATGGAAGAACTCGAAGCTCAGGA 1311
QY 461 GlnSerSerTyrProGlnArgThrSerAsnHisSerProProThrProArgArgSerPro 480
Db 1312 CAAAGTTCTTATCCCAAGGACCACAATCATAGTCTCTCTACCTCCGAGAGCCCT 1371
QY 481 IleProLeuAspArgProAspMetArgArgAlaAspSerLeuArgLysGlnHisLeu 500
Db 1372 ATACCCCTCGATAGACGAGCTGAGCGCTGCCGACTTCCCTACGGAACACACCACTTA 1431
QY 501 AspProSerSerAlaValArgLysThrLysArgLysMetGluThrMetLeuArgAsn 520
Db 1432 GATCCAGCTCTCTCTGTAAAGAAAACGAGCGAGAAAATGGAACAATGTTGAGGAAT 1491
QY 521 AspSerLeuSerSerAspGlnSerGluSerValArgProProProProArgProHisLys 540
Db 1492 GATTTCTTTGAGTTTCAGACCATCTCTGAGTCAGTGAGCGCGCCCTCACCAAGCCTCATAA 1551
QY 541 SerLysLysGlyLysMetArgGlnValSerLeuSerSerSerSerGluGluLeuAla 560
Db 1552 TCCAAGAAAGAGAGTAAATGCCAGTTTCTACAGCAGCTCAGAGGAGGAGTTGGCA 1611
QY 561 SerThrProGluTyrThrSerCysAspValGluLeuGluSerGluSerValSerGlu 580
Db 1612 TCCAGCCTGAGTATACAGCTGTGATGAGTGGAGATTGAAAGCGAGAGCCTAGGTGAG 1671
QY 581 LysGlyAspSerGlnLysGlyLysArgLysThrSerGluGlnGlyValLeuSerAspSer 600
Db 1672 AAAGGACATCGAG----- 1686
QY 601 AsnThrArgSerGluArgGlnLysLysArgMetTyrTyrGlyGlyHisSerLeuGluGlu 620
Db 1687 -----TACAGCTCGTGGAG 1701
QY 621 AspLeuGluThrSerGluProGlnLysLysAspSerGlyValAspThrCysSerSerThr 640
Db 1702 CATGCGTCTTGGCAT----- 1716
QY 641 ThrLeuAsnGluGluHisSerHisSerAspLysHisProValThrTrpGlnProSerLys 660
Db 1717 -----AGCAGTGGAGCATCCCCAATGCTTTTCACCTCTGACCTGGCAGCATCCAA 1770
QY 661 AspGlyAspArgLeuIleGlyArgIleLeuLeuAsnLysArgLeuLysAspGlySerVal 680
Db 1771 GATGGAGATCGCCTAAATGCTGCTATTTTATTAATAAGCGTTTAAAGATGGAGTGTA 1830
QY 681 ProArgAspSerGlyAlaMetLeuGlyLeuValValGlyLysMetThrGluSer 700

Db	1831	CCTCGAGATT	CAGGACCAAT	CTGGGCTTAAAGGTT	GAGGAGAAAGATG	CACTGAATCA	1890
Qy	701	GlyArgLeu	CysAlaPhe	LeuThr	ValVal	ValVal	ValVal
Db	1891	GGTCGACT	TTGTGCTAT	TATTTACAA	AGTGAAGAAAG	AGTGTAGCT	GATAGTAGGA
Qy	721	HisLeuArg	ProGly	AspGlu	ValLeu	GluThr	ProGlu
Db	1951	CACTTAGAC	CAAGGTGAT	GAAGCTT	CGAATGGA	ATGGAAGCT	TATTCGAGGAGCCACA
Qy	741	PheGluGlu	ValVal	ValVal	ValVal	ValVal	ValVal
Db	2011	TTTGAGGAA	GTATTACAA	CAATATTCT	AGAAATCA	AGCTCA	CAAGTTGAGCTTGT
Qy	761	ValSerArg	ProIle	GlyAsp	LeuPro	ArgLeu	ValLeu
Db	2071	GTTCAGGCG	GAATGGAGAT	TGAGATAC	CTGATAGCA	CTATGACCA	CACTGAA
Qy	781	SerSerSer	SerPhe	GluSer	GlnVal	MetAsp	ArgPro
Db	2131	TCCAGTTCT	AGCTCAT	TATTTGA	CTCAAAAT	TGATCGT	CGTATATCGTTACCTCT
Qy	801	ProMetSer	ProGly	MetLeu	ArgVal	ProGln	PheLeu
Db	2191	CCCATGAGT	CTCTGCGAT	CTGAGGAT	GTCCCA	CAGTTCT	TATCTGACACAGCTTCAATA
Qy	821	LysLeuTrp	PheAsp	ValVal	HisGln	LeuVal	ValVal
Db	2251	AAACTATG	TTGGTGA	CAAGGTTG	GTCCAGT	TATATAGT	TCAATTTTGGGAGCAAGGAT
Qy	841	LeuProSer	ArgGlu	AspGlu	ValPro	ArgAsn	ProVal
Db	2311	CTCCCTTCC	AGGAGAT	GGGAGCC	CAAGGAT	CCCTTAT	GTAAATTTACTTCTTCCA
Qy	861	AspArgSer	AspLys	AsnLys	ArgThr	LysThr	ValLys
Db	2371	GACAGAAGT	GTATAAAA	CAAGAGAA	CAACAA	CAGTCA	AGAACTTTGGACCCAAA
Qy	881	TrpAsnGln	ThrPhe	IleVal	SerPro	ValHis	ArgGlu
Db	2431	TGGAACCA	GACTTTT	CATTAT	TTCTCCT	GTCCAGG	AGAAATTC
Qy	901	GluIleThr	LeuTrp	AspGln	AlaArg	ValArg	GluGlu
Db	2491	GAAATCAC	CCCTTTT	GGGATCA	AGCTTC	GAGTTC	GAGGAGAAATTCCTAGGAG
Qy	921	IleLeuIle	GluLeu	GluThr	AlaLeu	AspAsp	GluPro
Db	2551	ATTTTAA	TGAAAT	TGAAAC	CAGCTCT	GTAGAT	GTACACCACTGGTACAACTTCAG
Qy	941	ThrHisAsp	ValSer	LeuPro	LeuPro	ArgPro	SerPro
Db	2611	ACCCATGAT	GTCTCCT	CATTGCCA	CTCCCT	CACCTTC	CCATATATGCCACGAGACAG
Qy	961	LeuHisGlu	LeuSer	ProThr	ArgArg	GlnArg	SerLys
Db	2671	CTGCATG	AGAGAGC	CCCAAC	AGGTTG	CAAGGTC	GAAGAAATAGTACAGCGAA
Qy	981	ValSerAsp	TrpAsp	CysGlu	AspGly	ValGly	ValVal
Db	2731	GTATCCG	ACTACT	GACTCG	GAGCGG	CTGGAGT	CTGTGTCAGATACCCACACATGGC
Qy	1001	ArgAspLeu	GlnSer	SerThr	LeuSer	ValPro	GluGln
Db	2791	CGTGATCTT	CAAGCTC	CAATAT	TATCAGT	GCAGAA	CAAGTCATCAAAACCAATGC
Qy	1021	SerProSer	GlySer	ProHis	ArgVal	AspVal	IleGly
Db	2851	TCACCATC	AGGCTCT	CTCAT	TCGATAG	ATGTTAT	TAGGAAGGACTAGTGTCTGCTGCT
Qy	1041	SerAlaPro	ProGln	ArgAsn	ValGlu	GlnGly	HisArg
Db	2911	AGTGTCC	CTCTCT	CTCAAG	AAATGT	GGACAG	GGGCTTCCAGGAGACACGTGCTACTGGC
Qy	1061	HisTrpAsn	ThrIle	SerArg	MetAsp	ArgHis	ArgVal
Db	2971	CATTACAA	TACATTA	TAGCCG	GTGACAC	ACCGTGT	CATGACCACTACTTCTTCA
Qy	1081	AspArgAsp	ArgCys	GluAla	AlaAsp	ArgGln	ProTrp
Db	3031	GAGAGAC	AGCAGT	CATTTCT	CACTCT	CACTCT	CGACACAGGACAGCAGTACGAT
Qy	1101	GluGluArg	ProLeu	LeuGlu	ArgThr	ThrArg	SerSer
Db	3091	CACCACAG	GGAT				
Qy	1121	ThrAsnLeu	MetArg	SerMet	ProSer	LeuMet	ThrGly
Db	3102						
Qy	1141	AlaLeuSer	ArgSer	HisPro	ArgThr	GlySer	ValGln
Db	3103						
Qy	1161	GlyThrGly	ArgArg	GlyArg	GlnLeu	ProGln	LeuPro
Db	3157	GTAAACAG	ACGAA	AGGCGG	CGGACG	TTCTC	CACTCCACAAAGGAGACACTGGAAGA
Qy	1181	SerAlaMet	AspIle	GluArg	AsnArg	GlnMet	LysLeu
Db	3216						
Qy	1201	AlaGlySer	AspPro	ArgLeu	GluGln	AspTrp	HisSer
Db	3216						
Qy	1221	ProHisArg	GlyAla	AspThr	ValSer	ThrLys	SerSer
Db	3216						
Qy	1241	SerAlaVal	SerArg	ThrSer	SerAla	SerArg	PheSer
Db	3216						
Qy	1261	GlnSerGlu	ArgPro	ArgGly	AsnArg	ValIle	SerVal
Db	3216						
Qy	1281	ArgGlnMet	GlyVal	SerGly	LysAsn	LeuThr	LysSer
Db	3216						
Qy	1301	CysSerLeu	GluLys	AsnAsp	GlySer	GlnSer	AspThr
Db	3216						
Qy	1321	SerGlyLys	LysArg	SerSer	IleGly	AlaLys	MetVal
Db	3216						
Qy	1341	ArgLysSer	ArgSer	AlaSer	GlnLeu	SerGln	ThrGlu
Db	3217						
Qy	1361	SerThrVal	GlnArg	SerThr	GluThr	GlyLeu	AlaVal
Db	3238	AGCAGCT	GTCCAG	AGAGAC	CGGAGC	CGGCTAG	CAGTGGAGATGAGAACTCGATGACA
Qy	1381	ArgGlnAla	SerArg	GluSer	ThrAsp	GlySer	MetAsn
Db	3298	AGACAGCC	ACGCGG	GAATCC	ACGATGG	CAGATG	CAACAGTTACAGCTCAGAGGAAAT
Qy	1401	LeuIlePhe	ProGly	ValArg	LeuAla	SerAsp	SerGln
Db	3358	CTGATCTT	CTCTGG	GGTCCG	CTCTG	ACGAC	CGAGTTCAGCGATTTCTTGTGGT

Qy 1421 LeuGlyProAlaGlnLeuValGlyArgGlnThrLeuAlaThrProAlaMetGlyAspIle 1440
Db 3418 CTGGCCCTGCTCAGCTAGTGGGACGCCAGACCTCGCGACTCTCGCAATGGTGACATT 3477
Qy 1441 GlnValGlyMetMetAspLysGlyGlnLeuGluValGluIleIleArgAlaArgGly 1460
Db 3478 CAGGTGGGATGATGGATAGAGGGACAGCTGGAGGTAGAGATCATCCGGGCTCGAGGC 3537
Qy 1461 LeuValValLysProGlySerLysThrLeuProAlaProTyrValLysValTyrLeuLeu 1480
Db 3538 CTTGTAGTAAACACAGGTTCACAGACACTGCCAGCACCATATGTCAAGGTGTACCTGTTA 3597
Qy 1481 AspAsnGlyValCysIleAlaLysLysLysThrLysValAlaLysLysThrLeuGluPro 1500
Db 3598 GACAAACGAGTCTGCATAGCCCAAAAGAAACCAAGGTGGCGAGAGACCTGGAGCCC 3657
Qy 1501 LeuTyrGlnGlnLeuLeuSerPheGluGluSerProGlnGlyArgValLeuGlnIleIle 1520
Db 3658 CTGTATCAGCAGCTGTTATCTTTCGAGAGAGACCCCGAGGGGAGGTGTACAGATCATT 3717
Qy 1521 ValTrpGlyAspTyrGlyArgMetAspHisLysSerPheMetGlyValAlaGlnIleLeu 1540
Db 3718 GTCTGGGAGATTATGGCCGATGGATCACAAATCCTTTATGGAGTGGCCCGACATATC 3777
Qy 1541 LeuAspGluLeuGluLeuSerAsnMetValIleGlyTyrPheLysLeuPheProProSer 1560
Db 3778 TTAGATGAACCTGGAACTATCCAACATGGTAATTGGATGGTTCAAACTCTTCCCTCTTCC 3837
Qy 1561 SerLeuValAspProThrSerAlaProLeuThrArgArgAlaSerGlnSerSerLeuGlu 1580
Db 3838 TCCCTAGTAGATCCAACTTGGCACCTCTGCAAGAGAGGCTTCCCAATCGTCTCTGAA 3897
Qy 1581 SerSerThrGlyProSerTyrSerArgSer 1590
Db 3898 AGTTCCTAGGACCTTCTTACTCTCGTTCA 3927

Search completed: November 16, 2003, 13:14:57
Job time : 13295 secs